Loeper, A. (2017). Cross-border externalities and cooperation among representative democracies. European Economic Review, v. 91, pp. 180-208.

Available in: https://doi.org/10.1016/j.euroecorev.2016.10.003

Cross-border externalities and cooperation among representative democracies

Antoine Loeper

Universidad Carlos III de Madrid, Spain

JEL classification:
D62
D72
F55
H41
H77

Keywords:
International Cooperation
Bargaining
Federalism
Externalities
Public good
Election
Strategic delegation

A B S T R A C T

This paper analyzes the provision of public goods with cross border externalities by representa tive democracies. The level of provision of each country is decided by a policy maker elected by majority rule at the country level. We compare the case in which policy makers set their policies noncooperatively with the case in which they set their policies through Coasian cooperation. Cooperation induces policy makers to internalize cross border externalities, but it also induces strategic voters to elect a policy maker who cares less about the public good to reduce their public good contribution. The former effect increases public good provision while the latter reduces it. We show that once voters' incentives are taken into account, whether cooperation is beneficial depends neither on voters' preferences, nor on the magnitude of spillovers, nor on the size, bargaining power, or efficiency of each country. Instead, it depends only on the curvature of the demand for the public good: cooperation increases (decreases) public good provision when the demand function is more (less) convex than the unit elastic demand function. Hence, the desirability of international cooperation depends mostly on the type of public good considered.

1. Introduction

Cross border externalities and transnational public goods lead to inefficiencies and collective action failure when countries set their policies noncooperatively. In the absence of overarching political institutions, observers often call for greater coordination between national policy makers to internalize these externalities. However, despite the multiplication of international negotiations and summits, the supposed gains from international cooperation have arguably not fully materialized. Many global public goods such as the reduction of greenhouse gas emission, political asylum, disease eradication, fish stocks, or fiscal stimulus are still underprovided.

Since Coase (1960), economists have invoked transaction costs of various sorts to explain the inability of bargaining parties to reach mutually beneficial arrangements.[1] This strand of literature focuses on the bargaining process and does not take into account the specificities of the political process within each country. Others have argued that international policy coordination can exacerbate inefficiencies in national politics.[2] This paper assumes away any inefficiencies in national politics or in Coasian cooperation, and focuses instead on the *interaction* between elections at the national level and cooperation at the international level.

In modern democracies, most decision are taken not by the voters, but by political representatives appointed by the voters. As Persson and Tabellini (1992) first pointed out, even if one abstracts away from political agency issues, this distinction has important

E-mail address: aloeper@eco.uc3m.es.

[1] Among others, commitment and enforcement problems (Williamson, 1985; North, 1990; Acemoglu, 2003) or imperfect information (Mailath and Postlewaite, 1990; Harstad, 2007) can lead to inefficiencies.

[2] International policy coordination can exacerbate political agency problems (Brennan and Buchanan, 1980; Buchanan and Faith, 1987; Tabellini, 1990; Persson and Tabellini, 1995) or dynamic commitment problems between voters and politicians (Rogoff, 1985; Kehoe, 1989).

consequences, because sophisticated voters can use elections as a strategic delegation mechanism. Several papers have shown that once voters' incentives are taken into account, the impact of international cooperation on public good provision is ambiguous (Segendorff, 1998; Gradstein, 2004; Buchholz et al., 2005; Kempf and Rossignol, 2013). On the one hand, cooperation helps national policy makers internalize cross border spillovers. This direct effect increases public good provision. On the other hand, more public good requires greater contributions from participating countries. As a result, cooperation induces strategic voters to elect representatives who care less about the public good, so as to decrease their relative contribution. This electoral effect decreases public good provision. In this paper, we determine the main drivers of the magnitude of this electoral effect, and characterize the conditions under which it only mitigates, or completely offsets the direct effect of cooperation.

We consider a model with two countries populated by a continuum of heterogeneous voters. Each country must decide the level of provision of a public good with cross border externalities. The preferences of a given voter are characterized by a type that determines her trade off between public good and private good consumption. This type can be interpreted as her *tax price* for the public good, and the mapping between her tax price and her most preferred level of public good is the *public good demand function*. Countries' policies are determined in a two stage game. In the first stage the electoral stage each country elects a representative among its residents by majority rule. In the second stage the policy making stage the elected representatives choose, cooperatively or noncooperatively, the level of provision of their respective public good. In the cooperative regime, the representatives implement the generalized Nash bargaining solution with the noncooperative outcome as the bargaining default.

The main result is that whether cooperation increases the equilibrium level of public good relative to the noncooperative regime depends neither on the distribution of voters' preferences, nor on the magnitude of cross border spillovers, nor on the relative size, efficiency, or bargaining power of each country. Instead, it depends only on the curvature of the demand for the public good. In the basic model, cooperation increases (decreases) public good provision if the public good demand function is more (less) convex than the unit elastic demand. This result holds unchanged for a large class of bargaining solutions. Making the two public goods closer substitutes makes cooperation more likely to be beneficial, but does not change the qualitative nature of the result. The model further shows that once voters' incentives are taken into account, allowing for transfers across countries can make cooperation detrimental.

That the desirability of cooperation is independent of the magnitude of spillovers may appear surprising, because the magnitude of the spillovers determines the inefficiency of the noncooperative equilibrium, and thus the potential gains from cooperation. The intuition for that result is as follows. Relative to the noncooperative equilibrium, cooperation requires the policy maker of, say, country 1 to provide more public good, and thus its voters to pay higher taxes. This impact of cooperation on the voters of country 1 is due to the internalization of the externality they impose on country 2. Therefore, as the magnitude of this externality increases, the cost of cooperation on the voters of country 1 increases. In the electoral stage, these voters react to the greater cost of cooperation by appointing a representative with a higher tax price for the public good, thereby offsetting the effect of greater spillovers in the policy making stage of the cooperative regime.

The intuition behind the role of the convexity of the public good demand function is more subtle and stems from the distributive effect of Coasian cooperation. As first intuited by Schelling (1960), since Coasian cooperation tends to equalize the gains from cooperation, it generates incentives to strategically delegate the negotiations to an agent who has less to gain from cooperation. Whether these incentives induce voters to elect a higher or a lower tax price representative turns out to depend solely on the curvature of the demand function. The reason for this is as follows. Cooperation basically prescribes the representative of each country to behave as if his tax price for the public good is subsidized at a rate that corrects for the externality, even though the subsidy is not actually paid. Thus, the cost of cooperation for each representative is the deadweight loss of a proportional subsidy. Simple calculus shows that the deadweight loss of a proportional subsidy, as given by the well known Harberger formula, increases (decreases) in the pre subsidy price when the demand function is less (more) convex than a unit elastic demand. So when the public good demand is, say, less convex than a unit elastic demand, as the tax price of the representative increases, his private cost of internalizing the externality its country imposes increases. Coasian cooperation compensates him for this greater cost by tilting the bargaining outcome in his favor and requiring a greater contribution from the other country, relative to the case of a unit elastic demand function. This distributive effect of cooperation in the policy making stage exacerbates voters' incentive to appoint a higher tax price representative in the electoral stage, which decreases public good provision.

Our results shed light on the debate over the structure of federal systems. Several competing principles have been invoked to determine the optimal allocation of policy responsibilities between central and local governments. One of them, “cooperative federalism,” states that federal policies should be negotiated by and “agreed to unanimously by the elected representatives from each of the lower tier governments” (Inman and Rubinfeld, 1997). Our analysis suggests that strategic voting can greatly affect the supposed gains from cooperative federalism. Moreover, contrary to received wisdom (Oates, 1972), whether cooperative federalism dominates decentralization depends neither on the heterogeneity of local preferences nor on the magnitude of externalities, but on the curvature of the demand for the public good, and thus on the type of public good.

Our results can be related to the empirical literature on the demand for public goods. This literature typically assumes isoelastic demand functions. For such demand functions, our results imply that cooperation is beneficial if and only if the tax price elasticity of the public good demand is greater than 1. Interestingly, empirical estimates of this elasticity vary greatly between public goods (see, e.g., Feldstein, 1975; Brooks 2007). Hence, our results imply that the efficiency of interjurisdictional cooperation can differ importantly across types of public goods. Moreover, estimated elasticities are often smaller than one (Wildasin, 1987; Auten et al., 2002). Therefore, this model suggests that for plausible specifications, strategic voting can severely offset the gains from Coasian cooperation between sovereign democracies. Hence, stronger forms of cooperation are required, such as pooling of sovereignty, or explicit cost sharing.

The paper is organized as follows. Section 2 reviews the literature. Section 3 describes the basic good model. Section 4 derives the main results. Section 5 considers alternative bargaining solutions. Section 6 allows for transfers across countries. Section 7 considers an alternative public good model with crowding out effects. Section 8 relates the results to the empirical literature on the demand for public goods.

2. Related literature

Since Olson (1965), the literature on collective action and public good provision has shown that the inefficiency of noncooperative behavior is more severe and thus the potential gains from cooperation are greater when spillovers are large (Oates, 1972; Sandler, 1998), or when preferences are homogeneous (Cornes, 1993). In contrast, in our model, the distribution of spillovers and preferences do not matter. Our results differ because the aforementioned literature focuses on the coordination failure between policy makers whereas this paper assumes that policy makers cooperate efficiently, and focuses instead on the coordination failure between voters of different jurisdictions.

A number of papers have shown that strategic delegation via elections can distort the outcome of Coasian cooperation between democratic jurisdictions. Brückner (2000) shows how to restore efficiency by appropriately allocating the proposal power. Harstad (2008) focuses on the impact of transfers on electoral incentives. Kempf and Rossignol (2013) investigate how equity considerations affect the success of international agreements, but do not compare electoral equilibria with and without cooperation. In Section 5, we show that once voters' incentives are taken into account, whether cooperation is beneficial or not does not depend on how egalitarian bargaining is. Most closely related to this paper, Segendorff (1998), Gradstein (2004), and Buchholz et al. (2005) also investigate whether strategic delegation offsets the gains from international cooperation. Segendorff (1998) and Gradstein (2004) consider a symmetric, pure public good model, and a unilateral externality model with transfers, respectively. Our results show that their conclusion are driven by their particular specifications for the cost structure of the public good, which determine the curvature of their demand functions. Buchholz et al. (2005) consider a symmetric environment with transfers, and do not fully characterize the determinants of the desirability of cooperation.[3]

This paper departs from the aforementioned contributions in that we focus on the empirically more relevant case in which international cooperation is carried out without monetary transfers, and political leaders stay in place in case of negotiation failure. More importantly, we consider a large class of public good environments and bargaining solutions, and characterize the conditions under which strategic voting makes cooperation detrimental. Our analysis singles out the curvature of the demand for the public good as the main driver of the cost of strategic voting. Even though this parameter has received a lot of attention in the empirical literature (see Section 8), its central role has been overlooked by the theoretical literature on strategic delegation.

Some papers analyze the impact of strategic delegation on more institutionalized forms of cooperation (Chari et al., 1997; Cheikbossian, 2000; Besley and Coate, 2003; Redoano and Scharf, 2004; Dur and Roelfsema, 2005; Harstad, 2010; Christiansen, 2013). These models differ from ours in that cooperation is carried out through federal institutions, fiscal arrangements, and/or majoritarian decision making. In contrast to the case of voluntary bargaining, these forms of cooperation generate electoral incentives that lead to overprovision. Common fiscal pools induce voters to elect public good lovers to increase central spending in their preferred public good. Majoritarian decision making leads to expropriation by the winning coalition, and thus induces voters to elect representatives that are biased in favor of the public project so as to be included in the winning coalition.

3. The basic model

We consider an economy composed of two countries (which can also be viewed as members of a federation). Throughout, $c \in \{1, 2\}$ refers to an arbitrary country and c to the other country. Each country c is inhabited by a continuum of voters I_c, and for all $i \in I_c$, (i, c) refers to voter i in country c. The voters of country c must choose the level of provision $x_c \geq 0$ of a national public good which can also be consumed to some degree by the residents of the other country, and which is financed by the taxes of the voters of country c. If $p_c > 0$ denotes the unit price of the public good and $s_{i,c} > 0$ the tax share of a voter (i, c) which can vary across voters because of income inequality then a level of provision x_c requires a tax contribution $p_c s_{i,c} x_c$ from voter (i, c).

3.1. Voters' preferences

The policy preferences of voter (i, c) are derived from the following utility function:

$$U_{i,c}(x) = t_{i,c}(G_c(x_c) + \beta_{-c} G_{-c}(x_{-c})) - s_{i,c} p_c x_c. \tag{1}$$

The term $t_{i,c}(G_c(x_c) + \beta_{-c} G_{-c}(x_{-c}))$ in (1) corresponds to the utility that voter (i, c) derives from the consumption of the domestic and the foreign public goods, whereas the term $-s_{i,c} p_c x_c$ corresponds to the foregone consumption of private goods due to the tax contribution needed to finance x_c. Thus, $t_{i,c} > 0$ captures how voter (i, c) trades off public good versus private goods consumption, and the parameter $\beta_{-c} > 0$ captures the magnitude of the externality coming from the foreign public good.

Throughout, we assume that for all $c \in \{1, 2\}$, G_c is twice continuously differentiable on $(0, +\infty)$, $G'c > 0$, $G''c < 0$, $\lim_{x\to 0} G'c(x) = +\infty$, and $\lim_{x\to\infty} G'c(x) = 0$.

[3] The relation between this paper and Segendorff (1998) is discussed in Section 4.3. The relation with Gradstein (2004) and Buchholz et al. (2005) is discussed in more detail in the working paper version (Loeper, 2015).

Some remarks about the specification (1) are in order. This specification imposes minimal restrictions on the public good technologies G_1 and G_2. As we shall see, this degree of freedom turns out to be crucial in our model. It also allows for any degree of preferences heterogeneity within and across countries via the distributions of the preferences and tax parameters $(s_{i,c})_{i\in I_c}$, $(t_{i,c})_{i\in I_c}$, and the public good prices (p_1, p_2). Note that these parameters can capture differences across countries in population sizes, or in taxation/public good provision efficiency. This specification also allows for spillovers of arbitrary magnitude and asymmetries, via (β_1, β_2). Finally, note that x_1 and x_2 can be public as well as private goods. Their only distinctive features is that they are publicly provided within each country, and that they spill over to the other country. However, to fix ideas, in what follows, we refer to them as public goods.

The specification (1) makes two main simplifying assumptions. First, voters have separable preferences over the domestic and the foreign public good, which means that voters' willingness to pay for their own public good is independent of the contribution of the other country. We consider an alternative public good model that relaxes this assumption in Section 7, and show that the main results are qualitatively unchanged. Second, the term $-s_{i,c}p_c x_c$ in (1) implicitly assumes that preferences are quasi linear in after tax income. This assumption can be viewed as a first order approximation which is justified if the cost of provision of the public good under consideration is not a large share of the total budget of each country. In the working paper version (Loeper, 2015), we show that our results can be easily adapted to allow for tax distortions that increase with the level of taxation, and for preferences that are concave in after tax income.[4]

Finally, note that the separability and quasi linearity assumptions are widely used in the literature on strategic delegation in public good settings, and the specification (1) encompasses that of most models in this literature, which facilitates the comparison of their results with ours.[5]

3.2. Individual tax prices and the demand for the public good

The utility function of voter (i, c) in (1) can be rescaled as follows:

$$U_c(p_{i,c}, x) = G_c(x_c) + \beta_{-c}G_{-c}(x_{-c}) - p_{i,c}x_c, \tag{2}$$

where $p_{i,c} \equiv s_{i,c}p_c/t_{i,c}$. Thus, the type $p_{i,c}$ completely characterizes the preferences of voter (i, c). The type of the representative and median voter of each country c are denoted by p_c^r and p_c^m, respectively, and we refer to a representative as "he" and to a voter as "she".

If voter (i, c) could choose the policy of her own country x_c taking the policy of the other country x_{-c} as given, the resulting level of public good would be $(G'c)^{-1}(p_{i,c})$. Thus, one can interpret $p_{i,c}$ as her *tax price* for the public good, and the function $D_c \equiv (G'c)^{-1}$ as the *public good demand function*. Its tax price elasticity is defined as usual as

$$\varepsilon_c(p) \equiv \frac{|D'c(p)|p}{D_c(p)}.$$

To illustrate our results, we occasionally use the following specification for G_c: for all $x_c > 0$ and $\varepsilon_c > 0$,

$$G_c(x_c) \equiv \frac{\varepsilon_c}{\varepsilon_c - 1}(x_c)^{\frac{\varepsilon_c - 1}{\varepsilon_c}}. \tag{3}$$

The public good demand induced by (3) is $D_c(p) = p^{-\varepsilon_c}$. Hence, its tax price elasticity is constant and equal to ε_c. If we let $\varepsilon_c \to 1$ in (3), $G_c(x_c) \to \ln(x_c)$, and the corresponding demand is $D_c(p) = 1/p$, which we refer to as the *unit elastic demand function*.

Definition 1. A demand function D is more (less) convex than another demand function $\widetilde{D}$ if $D = f(\widetilde{D})$ for some strictly convex (concave) function f.

The order "more convex than" can be viewed as a generalization of "more elastic than" in the sense that they coincide on isoelastic demand functions: if D and $\widetilde{D}$ have a constant elasticity of ε and $\tilde{\varepsilon}$, respectively, D is more convex than $\widetilde{D}$ if and only if $\varepsilon > \tilde{\varepsilon}$.

Since the public good demand function and its curvature play a central role in our analysis, it is worth discussing their economic meaning. Note that $p_{i,c}$ is not stricto censu a tax price, as it combines an actual tax price $s_{i,c}p_c$ and a preference parameter $t_{i,c}$. Economists typically define and estimate public good demand functions as functions of the unit price p_c or of the actual tax price $s_{i,c}p_c$. However, in this model, defining the public good demand of an individual as a function of p_c or $s_{i,c}p_c$ instead of $p_{i,c}$ would amount to a change of price unit. Such changes affect neither the elasticity nor the convexity of the demand curve.

3.3. Elections

We model the strategic interactions among voters and policy makers as a two stage game. In the first stage *the electoral stage* the voters of each country elect their respective representative, taking the representative of the other country as given. In the second

[4] Note that the term $-s_{i,c}p_c x_c$ also implicitly assumes that the unit price of the public good p_c is independent of level of provision. This assumption is without loss of generality, as one can always measure the level of public good x_c in terms of total cost of provision, change G_c accordingly, and normalize p_c to 1.

[5] To the best of our knowledge, all the papers on strategic delegation in public good settings assume quasi-linear preferences. Most of them also assume separable preferences. For instance Segendorff (1998), Cheikbossian (2000); Besley and Coate (2003), Gradstein (2004), Dur and Roelfsema (2005), or Kempf and Rossignol (2013) use special cases of our specification. An exception is Buchholz et al. (2005) who consider a public bad version of the model analyzed in Section 7.

stage *the policy making stage* elected representatives set the policy of their respective country, cooperatively or noncooperatively.

To focus on the role of elections as delegation mechanisms and abstract away from other electoral effects such as electoral competition or political agency, we model national elections via a simplified "citizen candidate" model as in Persson and Tabellini (1992). Representatives are selected from the pool of voters, and any voter is willing to become the representative of her country if she receives enough votes. Candidates cannot make credible electoral promises (Alesina, 1988). Therefore, once elected, policy makers behave according to their own preferences.

Voters are consequentialist and rational: they care only about the policies and can foresee how the type of their representative affect the policy outcome. Formally, for any profile of representatives' types (p_1^r, p_2^r) elected in the electoral stage, if $x^{PM}(p_1^r, p_2^r)$ denotes the corresponding outcome in the policy making stage, then the *induced preferences* of voter (i, c) on her representative $p_c{}^r$ is given by $p_c^r \to U_c(p_{i,c}, x^{PM}(p_1^r, p_2^r))$.

To abstract away from the issue of voters' miscoordination, we assume that the election winner in a given country is the citizen who is preferred by a majority of voters to any other candidate.[6] The following proposition states that in our environment, the national median type will always be pivotal in the electoral stage.

Lemma 1. *For all $c \in \{1, 2\}$, for any pair of policy vectors $x, x' \in \mathbb{R}_+^2$, a majority of voters in country c prefer x to x′ if and only if the median type $p_c{}^m$ prefers x to x′.*

Therefore, in the electoral stage, irrespective of the outcome of the policy making stage, the induced majority preferences of the voters of country c on their representative $p_c{}^r$ always coincide with the induced preferences of $p_c{}^m$.

Depending on the nature of the public good, voting on the type of the representative can be interpreted as choosing a candidate with a particular belief about the intrinsic value of the public good (e.g., a global warming skeptic, a monetarist in the case of fiscal stimulus, or an anti immigration politician in the case of political asylum), a special inclination towards the polluting industry or the environmental lobbies, or a candidate who seek the support of voters from a particular income group, and thus from voters with a particular tax price.

4. Public good provision by elected representatives

4.1. The noncooperative regime

We first analyze the benchmark case in which policy makers behave noncooperatively. Formally, we assume that in the policy making stage, for any profile of representatives p^r appointed in the electoral stage, the elected representative of each country c chooses the policy x_c that maximizes his welfare, taking the policy of the other country x_{-c} as given. The unique dominant strategy for each representative is $x_c = D_c(p_c^r)$, so the outcome of the policy making stage game is:

$$x_c^N(p^r) = D_c(p_c^r). \tag{4}$$

One can see from (4) that this policy outcome is independent of β: in the absence of cooperation, policy makers fail to internalize the cross border externalities.

Given the outcome of the policy making stage $x^N(.)$, the equilibrium of the electoral stage is defined by backward induction as follows:

Definition 2. A representative's type $p_c{}^r$ is a noncooperative electoral best response of the voters of country c to some p_{-c}^r if for all $p_c > 0$, $x^N(p_c^r, p_{-c}^r)$ is majority preferred in country c to $x^N(p_c, p_{-c}^r)$. A profile of type p^r is a noncooperative electoral equilibrium (henceforth NEE) if for all $c \in \{1, 2\}$, $p_c{}^r$ is a noncooperative electoral best response to p_{-c}^r.

Proposition 1. *The noncooperative electoral best response of the voters of country c to any $p_{-c}^r > 0$ is to elect their median voter p_c^m, and the corresponding level of public good in country c is $x_c^N(p_c^m) = D_c(p_c^m)$. Thus, the unique NEE is $p^r = p^m$. In the NEE, both public goods are underprovided: there exists a policy vector x that is strictly preferred to $x^{NEE} = x^N(p^m)$ by a majority of voters in both countries, and for any such x, $x_1 > x_1^{NEE}$ and $x_2 > x_2^{NEE}$.*

The intuition for Proposition 1 is as follows. Lemma 1 implies that the noncooperative electoral best response of country c to some p_{-c}^r is the type $p_c{}^r$ most preferred by its median voter p_c^m, given the policy outcome $p_c^r \to x^N(p_c^r, p_{-c}^r)$. From (4), $p_c{}^r$ determines $x_c{}^N$ but does not affect x_{-c}^N. Therefore, the delegation game in the electoral stage is strategically equivalent to a game in which each median voter $p_c{}^m$ chooses the policy of her country x_c, taking the policy of the other country x_{-c} as given. As a result, in the NEE, $p_c{}^m$ chooses her most preferred level of public good $x_c^{NEE} = D_c(p_c^m)$. The inefficiency of the NEE follows from the fact that median voters do not internalize externalities.

4.2. The cooperative regime

In view of Proposition 1, a natural question is whether Coasian cooperation between policy makers can mitigate the

[6] One can model elections as a normal-form game in which the strategy of each voter is the individual for whom she votes, and the outcome is the policy vector induced by the type of the plurality winner. This game typically has multiple Nash equilibria (e.g., all voters voting for a given type). However, whenever a type is a Condorcet winner, there is a strong Nash equilibrium in which a majority of voters vote for a candidate of that type, and all strong Nash equilibria are outcome equivalent, consistently with our assumption.

underprovision of the public good that arises in the NEE. To address this question, we modify the game analyzed in Section 4.1 by letting representatives cooperate in the policy making stage. Specifically, for a given profile of representative p^r elected in the electoral stage, we assume that the outcome of the policy making stage, which we denote $x^B(p^r)$, is the generalized Nash bargaining solution for the elected representatives. We set the bargaining default to be the equilibrium $x^N(p^r)$ that arises if representatives choose their respective policy noncooperatively, as characterized in (4).

Formally, for all policy vector $x \in \mathbb{R}_+^2$, let $\Delta_c(x, p^r)$ denote the payoff gain for representative $p_c{}^r$ from implementing x instead of the bargaining default $x^N(p^r)$. That is,

$$\Delta_c(p^r, x) \equiv U_c(p_c^r, x) - U_c(p_c^r, x^N(p^r)). \tag{5}$$

Then $x^B(p^r)$ is the policy vector that solves

$$\max_{x \in \mathbb{R}_+^2} B(\Delta_1(p^r, x), \Delta_2(p^r, x)), \tag{6}$$

where $B(\Delta_1, \Delta_2) \equiv \pi_1 \ln(\Delta_1) + \pi_2 \ln(\Delta_2)$, and $\pi_c > 0$ is the bargaining power of each country c. As shown in the Appendix A (see Lemma 2), the program (6) has a unique solution, and it satisfies

$$x_c^B(p^r) = D_c\left(\frac{p_c^r}{1 + \beta_c \frac{B_{-c}(p^r)}{B_c(p^r)}}\right), \tag{7}$$

where

$$B_c(p^r) \equiv \frac{\partial B_c}{\partial \Delta_c}(\Delta(p^r, x^B(p^r))). \tag{8}$$

In words, B_c is the utilitarian weight that the bargaining program (6) attaches to country c at its solution $x^B(p^r)$. So the ratio B_{-c}/B_c in (7) is the rate at which the bargaining function B trades off the gains from cooperation between the two representatives. If we take this ratio as an exogenous parameter in the formula (7), then as B_{-c}/B_c increases, the policy vector x^B remains Pareto optimal for the two representatives, but $x_c{}^B$ increases whereas x_{-c}^B decreases, so the cooperative outcome becomes less favorable to $p_c{}^r$ and more favorable to p_{-c}^r. Thus, the ratio B_{-c}/B_c captures the *distributive effect* of cooperation.

A few remarks about the cooperative regime are in order. First, even though countries' representatives cooperate, both countries remain sovereign in that representatives are elected by their respective electorate, policies are financed at the national level, and the bargaining outcome is mutually beneficial for the elected representatives. Second, the Nash bargaining solution is arguably the most natural way to capture the unstructured and voluntary nature of Coasian cooperation between leaders of sovereign jurisdictions. We consider alternative bargaining solutions in Section 5. Third, transfers across countries are ruled out. We relax that assumption in Section 6. Finally, the bargaining default $x^N(p^r)$ implicitly assumes that the elected representatives p^r stay in power in case of negotiation breakdown. This assumption is consistent with the empirical observation that reelections are rarely held on the ground that representatives failed to reach an agreement.[7]

Given the outcome of the policy making stage $x^B(p^r)$, the equilibrium in the electoral stage is defined as follows:

Definition 3. A representative's type $p_c{}^r$ is a cooperative electoral best response of the voters of country c to some p_{-c}^r if for all $p_c > 0$, $x^B(p_c^r, p_{-c}^r)$ is majority preferred in country c to $x^B(p_c, p_{-c}^r)$. A profile of type p^r is a cooperative electoral equilibrium (henceforth CEE) if for all $c \in \{1, 2\}$, $p_c{}^r$ is a cooperative electoral best response to p_{-c}^r.

The comparison of (4) with (7) shows that cooperation prescribes policy makers to behave as if their tax price was $\frac{p_c^r}{1+\beta_c B_{-c}/B_c}$ instead of $p_c{}^r$. Hence, cooperation requires each representative to behave as if his public good was subsidized at a rate $\tau_c = \frac{\beta_c B_{-c}/B_c}{1+\beta_c B_{-c}/B_c}$. The externality parameter β_c in that rate corresponds to *efficiency effect* of cooperation that is, the internalization of externalities whereas the ratio B_{-c}/B_c corresponds to the *distributive effect* of cooperation. Since $\tau_c > 0$, $x_c^B(p^r) > x_c^N(p^r)$, which means that holding the profile of representatives p^r constant, cooperation unambiguously increases public good provision. However, the next proposition shows that once voters' incentives are taken into account, whether cooperation increases public good provision depends on whether the distributive effect of cooperation, as captured by the ratio B_{-c}/B_c, is tilted in favor of higher or lower tax price representatives.

Proposition 2. *For all $c \in \{1, 2\}$ and $p_{-c}^r > 0$, there exists a cooperative electoral best response for the voters of country c. For any profile of type p^r, if $p_r{}^c$ is a cooperative electoral best response to p_{-c}^r, the corresponding level of public good $x_c^B(p^r)$ in country c is strictly smaller (greater) than at the noncooperative best response of country c to any $\tilde{p}_{-c}^r$ if $\frac{\partial[B_{-c}/B_c]}{\partial p_c^r} < 0 \ (>0)$.*

Proposition 2 states the following intuitive result: if the distributive effect of cooperation becomes more favorable to country c as the tax price of its representative $p_c{}^r$ increases that is, if $\partial[B_{-c}/B_c]/\partial p_c^r < 0$ then the distributive effect of cooperation exacerbates voters' incentives to elect a high tax price representative, which decreases public good provision.

Proposition 1 further states that the sign of $\partial[B_{-c}/B_c]/\partial p_c^r$ is the *only* determinant of the comparison between the cooperative and noncooperative electoral equilibrium. The intuition behind that result is that the distributive effect determines whether at the

[7] See Segendorff (1998) and Gradstein (2004) for more on this issue.

electoral stage of the cooperative regime, a greater contribution by the voters of a given country increases or crowds out the contribution of the other country. To see why, note that as in the noncooperative regime, when choosing the tax price of her representative p_c^r, the voters of country c indirectly choose their country's contribution: x_c^B increases as they elect a representative with a lower tax price (see Lemma 4 in the Appendix A). However, in contrast to the noncooperative regime, when choosing their contribution x_c^B via the type of their representative p_c^r, they do not take the contribution of the other country x_{-c}^B as given. For instance, when $\partial[B_{-c}/B_c]/\partial p_c^r < 0$, we see from (7) that $\partial x_{-c}^B/\partial p_c^r > 0$. This means that as the voters of country c increase their contribution x_c^B (by decreasing p_c^r), they decrease the contribution of the other country x_{-c}^B, which is detrimental to them. This distributive effect of cooperation reduce their incentive to provide their public good, relative to the noncooperative regime.

4.3. When does cooperation increase public good provision?

The next proposition states our main result using the notion of convexity in Definition 1.

Proposition 3. *For any $c \in \{1, 2\}$, if D_c is more (less) convex than the unit elastic demand $1/p$, any CEE yields strictly more (less) public good in country c than the NEE irrespective of the other parameters of the model.*

The most striking result in Proposition 3 is that once voters' incentives are taken into account, whether cooperation increases public good provision depends neither on the magnitude nor on the asymmetries of the spillovers, as captured by β, nor on voters' preferences as captured by the distribution of $p_{i,c}$ within and across countries, nor on the allocation of bargaining power, as captured by π. This result may appear surprising, since these parameters are the main drivers of the impact of cooperation in the policy making stage.

To illustrate the intuition behind this result, consider for instance an increase in the spillovers from country c or a decrease in its bargaining power that is, an increase in β_c or a decrease in π_c. If we fix the profile of representatives p^r, one can see from (4) and (7) that these parameter changes do not affect the noncooperative policy equilibrium x^N, whereas they increase the provision of public good by country c in the cooperative policy outcome x^B. Thus, they increase the additional tax payment that cooperation requires from the voters of country c. The rational response of the voters of country c is to mitigate the greater cost of cooperation in the policy making stage by electing a higher tax price representative in the electoral stage. Hence, the impact of these parameter changes in the policy making stage of the cooperative regime is offset by voters' noncooperative behavior in the electoral stage. This intuition explains why cooperation can be detrimental even for arbitrarily severe externalities, and thus even when the potential gains from cooperation are arbitrarily large.

The second result stated by Proposition 3 is that whether cooperation increases public good provision depends only on the public good technology G_c, via the curvature of the public good demand function D_c: cooperation is beneficial only when D_c is sufficiently convex. The intuition for this result is more involved and is best explained through a sketch of the proof, which we provide at the end of this section.

Proposition 3 can be rephrased as a decentralization theorem in the spirit of Oates (1972). Suppose that countries 1 and 2 form a federal system, that the noncooperative regime is interpreted as decentralization, and that centralization takes the form of voluntary cooperation among the members of the federation ("cooperative federalism" , as coined by Inman and Rubinfeld (1997)). Then Proposition 3 states that in sharp contrast to Oates' decentralization theorem, the comparative advantage of centralization depends neither on the heterogeneity of local preferences, nor on the magnitude of externalities. Instead, it depends only on the curvature of the demand for the public good.[8]

Most empirical papers on public good demand consider parametrized families of demand functions (see the discussion in Section 8). Our results can be readily applied to any such parametrized family. For instance, in the canonical case of the isoelastic public good demand specified in (3), Proposition 3 implies the following.

Corollary 1. *For any $c \in \{1, 2\}$, if D_c has a constant tax price elasticity ε_c, then irrespective of the other parameters of the model, any CEE yields strictly more (less) public good in country c than the NEE if $\varepsilon_c > 1$ ($\varepsilon_c < 1$).*

Segendorff (1998, Proposition 3) considers the same model as ours, with $G_1(x) = G_2(x) = -\exp(-x)$, $\beta_1 = \beta_2 = 1$, and $\pi_1 = \pi_2$. In that case, $D_c(p) = \ln(1/p)$, so Definition 1 and Proposition 3 imply that public good provision is lower in any CEE than in the NEE, but that result is driven solely by the particular public good technology assumed in that paper.

We conclude this section with a sketch of the proof of Proposition 3. As shown in Proposition 2, the equilibrium impact of cooperation in country c depends on how the tax price of its representative p_c^r affects the distributive effect of cooperation, as captured by B_{-c}/B_c. To understand the role of the curvature of D_c, it is enough to focus on the effect of p_c^r on B_c, which is the utilitarian weight allocated to representative c by the bargaining function B. In the case of Nash bargaining, B_c is inversely proportional to his gains from cooperation. This relationship captures the need for Coasian cooperation to be mutually beneficial, and thus to share the gains from cooperation. Hence, to tilt the distributive effect of cooperation in their favor i.e., to increase the utilitarian weight of their representative voters have an incentive to choose the tax price of their representative so as to reduce his gains from cooperation. These electoral incentives formalize Schelling's insight (Schelling, 1960) that a successful bargaining

[8] Besley and Coate (2003) consider a model of federalism with strategic delegation in which the magnitude of spillovers affects the desirability of centralization, as in Oates' decentralization theorem. The key difference between this model and ours is that the cost of public good provision is shared by all districts. The common pool effect generated by this fiscal arrangement induces voters to elect public good lovers to increase central spending in their preferred public good. These incentives are socially less costly when local public goods have greater spillovers. See also Harstad (2007) or Loeper (2011) on alternative models in which Oates' decentralization no longer holds.

strategy is to delegate the negotiations to an agent who has less to lose in case of negotiation breakdown.

How can voters implement Schelling's strategy in our setting? Recall that as argued in Section 4.2, cooperation requires representative c to behave as if his tax price $p_c{}^r$ was subsidized at a rate $\tau_c = \frac{\beta_c B_{-c}/B_c}{1+\beta_c B_{-c}/B_c}$ even though the subsidy is not actually paid. Thus, his cost of cooperation is simply the deadweight loss of the "as if" subsidy τ_c, whereas his benefit from cooperation is the spillover effect of the "as if" subsidy τ c on the tax price of the other representative. If we assume for simplicity that τ_c and τ_{-c} are constant, $p_c{}^r$ affects the cost but not the benefit of cooperation for representative c. Therefore, electing a representative with smaller gains from cooperation means electing a representative whose deadweight loss of the "as if" subsidy τ_c is greater. How this deadweight loss varies with the tax price of the representative turns out to depend on the curvature of the demand function: the deadweight loss of a proportional subsidy τ_c, as given by the well known Harberger formula, increases (decreases) in the pre subsidy price $p_c{}^r$ when the demand function D_c is less (more) convex than the unit elastic demand. To see this formally, note that this deadweight loss equals

$$DWL \equiv G_c(D_c(p_c^r)) - p_c^r D_c(p_c^r) - [G_c(D_c((1-\tau_c)p_c^r)) - p_c^r D_c((1-\tau_c)p_c^r)] \qquad = -\int_0^{\tau_c} \frac{t}{(1-t)^2}((1-t)p_c^r)^2 D'c((1-t)p_c^r)dt.$$

The second equation, which follows from elementary calculus, shows that DWL increases (decreases) in $p_c{}^r$ when $p \to p^2 D'c(p)$ is decreasing (increasing), or equivalently when D_c is less (more) convex than $1/p$. Thus, the convexity of the demand for the public good matters because it determines whether a higher tax price representative has smaller gains from cooperation, and thus whether voters' incentives to tilt the distributive effect of cooperation in their favor induce them to appoint a higher tax price representatives.[9]

4.4. When is cooperation socially beneficial?

Proposition 1 shows that the inefficiency of the NEE is due to the underprovision of the public good. Therefore, when cooperation further decreases public good provision that is, when voters' demand function for the public good is not too convex it should be socially detrimental. The following proposition formalizes this intuition.

Proposition 4. *For any $c \in \{1, 2\}$, if D_c is less convex than the unit elastic demand, then a majority of voters in country c are strictly worse off in any CEE than in the NEE.*

The following proposition provides a partial converse to Proposition 4.

Proposition 5. *If D_1 and D_2 are more convex than the unit elastic demand, then in at least one country, a majority of voters is strictly better off in the CEE than in the NEE.*

If countries are further symmetric that is, $D_1 = D_2$, $\beta_1 = \beta_2$, $p_1^m = p_2^m$ and $\pi_1 = \pi_2$ then in both countries, a majority of voters are strictly better off in any symmetric CEE than in the NEE.

Thus, Propositions 3 5 show that whether we look at public good provision or voters' welfare, the equilibrium impact of cooperation depends primarily on the degree of convexity of voters' public good demand function.

Note that from Proposition 3, when both D_1 and D_2 are more convex than the unit elastic demand, cooperation increases public good provision in both countries, but even in that case, Proposition 5 does not guarantee that cooperation is beneficial for a majority of voters in both countries. The intuition for this result is that the effort that cooperation requires in terms of extra public good provision may not be shared sufficiently equally across countries. To understand the reason for this asymmetric impact of cooperation, note first that if median voters appointed themselves in the cooperative regime, Coasian cooperation would always benefit both of them, and from Lemma 1, it would benefit a majority of voters in both countries. Thus, the excessively unequal impact of cooperation must come from the differences in voters' incentives for strategic delegation. The intuition given in Section 4 suggests that the strength of these incentives is negatively related to the degree of convexity of their public good demand function. This suggests that in the electoral equilibrium, cooperation will require a relatively greater effort from the country whose public good technology G_c induces the more convex public good demand. Cooperation might thus be detrimental to the voters of this country even when it result in more public good in both countries. The following proposition formalizes this intuition in the case in which one country has a unit elastic demand.

Proposition 6. *Suppose D_c is unit elastic and D_{-c} is more convex than D_c. Then a majority of voters in country c are strictly better off in any CEE than in the NEE, but a majority of voters in country c are strictly worse off in any CEE than in the NEE.*[10]

---

[9] In the Appendix A (see the footnote at the proof of Lemma 5), we explain the connection between this sketch of the proof and the actual proof. Note that for brevity, we have considered only the effect of $p_c{}^r$ on the terms of cooperation B_{-c}/B_c through its effect on B_c. But a change in $p_c{}^r$ also affects B_{-c}/B_c through B_{-c}. However, it turns out that the latter effect always go in the same direction as the former effect. To see this, recall that B_{-c} is inversely proportional to the gains of cooperation of representative c, so $p_c{}^r$ affects these gains via the spillover effect of the "as if" subsidy τ_c on the tax price of representative c. This spillover effect is given by

$$S \equiv \beta_{-c}[G_{-c}(D_{-c}((1-\tau_{-c})p_{-c}^r)) - G_{-c}(D_{-c}(p_{-c}^r))]$$
$$= -\beta_{-c}\int_0^{\tau_c}\frac{1}{1-t}[((1-t)p_c^r)^2 D_{\prime c}((1-t)p_c^r)]dt,$$

where the second equality follows from elementary calculus. Thus the spillover effect S increases (decreases) in p^r when $p \to p^2 D'c(p)$ is decreasing (increasing). In this case, an increase in $p_c{}^r$ increases (decreases) not only the cost of cooperation for representative c but also the benefit of cooperation for representative c, and both effects affect the terms of cooperation in the same way.

[10] Note that when D_c is unit elastic, and D_{-c} is more convex than D_c, Proposition 3 imply that in any CEE, the contribution of country c is strictly greater than in the

Note that Proposition 6 holds irrespective of the degree of asymmetry in the externalities, in the bargaining power, and in voters' preferences across countries. Thus, Proposition 6 confirms the above intuition that once voters' incentives are taken into account, it is the asymmetry in the degree of curvature of D_1 and D_2 that drives the distributive impact of cooperation.

4.5. Equilibrium existence

Proposition 1 shows that a NEE always exists. From Proposition 2, a cooperative electoral best response always exists. The next proposition further shows that a CEE exists under standard specifications of the functions G_1 and G_2, as well as when spillovers that are not too large.[12]

Proposition 7. *For any p^m, β, and π, if G_1 and G_2 are given by the isoelastic specification (3) with $\varepsilon = 1$, $\varepsilon = 2$,* [11] *or with ε sufficiently large, then a CEE exists. For any G_1, G_2, p^m, and π, for β sufficiently small, a CEE exists.*

Note also that for the sake of clarity, our main results are stated as equilibrium properties, but we show in the Appendix A (see Propositions 15 17 that they can also be stated as properties of the cooperative electoral best response of a given country, which always exists from Proposition 2.

5. Alternatives bargaining solutions

The basic model uses Nash bargaining to model Coasian cooperation among policy makers. The choice of this bargaining solution is motivated by its desirable properties,[13] and its widespread use. Nevertheless, our results hold unchanged for a much larger class of bargaining function.[14] Specifically, the proofs only assume that the function $B(\Delta_1, \Delta_2)$ used in the bargaining program (6) is differentiable, strictly increasing in Δ_1 and Δ_2, and that for all $c \in \{1, 2\}$,

$$\Delta_c(p^r, x^B(p^r)) \geq 0, \tag{9}$$

and

$$\frac{\partial\left(\frac{\partial B}{\partial \Delta_{-c}} \Big/ \frac{\partial B}{\partial \Delta_c}\right)}{\partial \Delta_c} > 0. \tag{10}$$

That B is increasing in each coordinate is equivalent to assuming that the bargaining outcome is Pareto efficient. Assumption (9) captures the requirement that B must be such that the bargaining outcome $x^B(p^r)$ is mutually beneficial. Assumption (10) can be viewed as a property of diminishing marginal rate of substitution. This property has a natural distributive interpretation: the ratio $\frac{\partial B}{\partial \Delta_{-c}} / \frac{\partial B}{\partial \Delta_c}$ is the rate at which the function B trades off the gains from cooperation among the two representatives. So (10) requires that as the bargaining default of agent c increases, and thus as his gain from cooperation Δ_c decreases, the marginal rate of substitution $\frac{\partial B}{\partial \Delta_{-c}} / \frac{\partial B}{\partial \Delta_c}$ decreases, and thus becomes more favorable to him.[15]

To illustrate these assumptions, consider for instance the C.E.S. specification for the bargaining function B: for all $\pi_1, \pi_2 > 0$ and $\rho < 1$,

$$B^\rho(\Delta_1, \Delta_2) \equiv \rho^{-1}(\pi_1(\Delta_1)^\rho + \pi_2(\Delta_2)^\rho)^{1/\rho},$$

where $\pi_c > 0$ is the bargaining power of country c. One can easily show that B^ρ satisfies all of the above properties. The limit case

(*footnote continued*)
NEE, whereas the contribution of country c is the same in the CEE as in the NEE. However, it should be clear that by continuity, the same result applies if D_c is isoelastic with elasticity ε_c greater than but sufficiently close to 1. In that case, the contribution of both countries is strictly greater in the CEE than in the NEE, but one country would still have a majority of voters strictly worse off in the CEE than in the NEE.

[12] Even though we were unable to find parameters values in (1) for which a CEE does not exist, it is difficult to establish the existence of a CEE for arbitrary parameter values. The reason is as follows. A CEE is basically the Nash equilibrium of a delegation game between the two median voters. The usual fixed point theorems used to prove equilibrium existence require the best response of each median voter i.e., the set of representative's types $p_c{}^r$ most preferred by $p_c{}^m$, given the policy outcome $p^r \to x^B(p^r)$ to have a closed graph, and to be convex. The former property follows immediately from the maximum theorem, and from the continuity of the mapping $p_c^r \to x^B(p^r)$ and $x \to U_c(p_c^m, x)$. The latter property is harder to establish, because the bargaining program (6) that defines $x^B(p^r)$ typically does not have a closed form solution, so the induced preferences of the median voters on his representative's type are difficult to characterize. The proof of Proposition 7 basically shows that in all the cases in which we were able to solve (6) analytically or asymptotically, the best response of each median voter is unique, and thus convex.

[11] The isoelastic public good demand with $\varepsilon = 1$ and $\varepsilon = 2$ correspond to the public good technology $G(x) = \ln(x)$ and $G(x) = \sqrt{x}$, respectively, as assumed in Besley and Coate (2003) and Gradstein (2004).

[13] In particular, the scale invariance property of the generalized Nash bargaining solution implies that rescaling (1) as (2) is without loss of generality. Without the scale invariance property, our results might depend on the particular affine transformation of (1) we use as inputs for the bargaining solution.

[14] The only results that are derived specifically for the Nash bargaining solution are Proposition 7 on the existence of a CEE (but Proposition 8 below shows that a CEE also exists with utiliarian bargaining), and Propositions 9 and 10 on the role of transfers.

[15] Note also that $\frac{\partial B}{\partial \Delta_{-c}} / \frac{\partial B}{\partial \Delta_c}$ differs from the ratio B_{-c}/B_c introduced in Section 4.2 in that the latter is evaluated at $\Delta = \Delta(p^r, x^B(p^r))$, whereas $\frac{\partial B}{\partial \Delta_{-c}} / \frac{\partial B}{\partial \Delta_c}$ is evaluated at any $\Delta \in \mathbb{R}_+^2$, and is thus an intrinsic property of the bargaining function B. Assumption (10) on $\frac{\partial B}{\partial \Delta_{-c}} / \frac{\partial B}{\partial \Delta_c}$ captures Schelling's intuition (1960) that a player who has less to loose from a negotiation breakdown is able to negotiate a better deal. To see that (10) is a mild assumption, note that with transferable utility, if we reverse the inequality in (10), then the bargaining solution always allocate all of the surplus to one agent.

$\rho \to -\infty$ corresponds to egalitarian bargaining, the intermediate case $\rho \to 0$ to Nash bargaining, and the limit case $\rho \to 1$ to utilitarian bargaining. Thus, as ρ increases, B^{ρ} puts a greater weight on efficiency i.e., maximizing the total gains from cooperation relative to distributive concerns i.e., allocating the gains from cooperation in a mutually beneficial and equitable way. That our results hold unchanged for any such function B^{ρ} shows that the equilibrium impact of cooperation is unaffected by how the bargaining process trades off efficiency and distributive concerns, as long as it puts some positive weight on the latter.

This conclusion contrasts with the findings in Kempf and Rossignol (2013) that equity considerations may affect the feasibility of international agreements. The difference between our results and theirs comes from the fact that they compare cooperation and no cooperation holding the identity of the policy makers fixed. Thus, they abstract away from the impact of cooperation on voters' incentives, which is the main focus of this paper.

As argued at the end of Section 4.3, the role of the curvature of the public good demand in our results is driven by the distributive effect of Coasian cooperation among the representatives, as captured by Properties (9) and (10). One can formalize this intuition by considering the limit case of utilitarian bargaining. This bargaining solution maximizes the gains from cooperation irrespective of any distributive concerns. Formally, when $\rho = 1$, $\frac{\partial B^1}{\partial \Delta_{-c}} / \frac{\partial B^1}{\partial \Delta_c} = \pi_{-c}/\pi_c$, which is independent of Δ, so Property (10) is violated. Utilitarian bargaining also violates Property (9) in our model when p^r is sufficiently heterogeneous. In line with the above intuition, the following proposition shows that when these distributive properties are violated, the equilibrium impact of cooperation is still independent of voters' preferences and the magnitude of externalities, as under Nash bargaining, but it is also independent of the curvature of D_c.

Proposition 8. *Suppose representatives use the utilitarian bargaining solution in the policy making stage of the cooperative regime. Then there exists a unique CEE, and it leads to the same policy outcome as the NEE.*

6. Transfers

Side payments are usually beneficial in bargaining situations because they increase the potential gains from cooperation. In this section, we investigate whether they remain beneficial when negotiations are conducted by representatives elected by strategic voters. To do so, we modify the basic model of Section 3 and assume that each country can costlessly transfer some of its tax receipts to the other country. Formally, in the utility function (1) of an arbitrary voter (i, c), the term $s_{i,c} p_c x_c$ that captures her tax payment is replaced by $s_{i,c}(p_c x_c + t_c)$, where t_c denotes the transfer made by state c to the other state, with $t_1 + t_2 = 0$. As in the basic model, we rescale the utility function as follows

$$\begin{cases} V_{i,1}(x, t) = G_1(x_1) + \beta_2 G_2(x_2) - p_{i,1} x_1 + p_{i,1} t, \\ V_{i,2}(x, t) = G_2(x_2) + \beta_1 G_1(x_1) - p_{i,2} x_2 - \alpha p_{i,2} t, \end{cases}$$

where $t \equiv -t_1/p$, $\alpha \equiv p_1/p_2$ is the production efficiency of country 2 relative to country 1, and $p_{i,c}$ is as in the basic model. The NEE and CEE are defined as in Section 4 (see Definitions 2 and 3).

Clearly, no transfers are made in the policy making stage of the noncooperative regime, so the characterization of the NEE in Proposition 1 holds unchanged. In the cooperative regime, the outcome of the policy making stage is the solution to (6) with transfers, which yields

$$x^B(p^r) = \left(D_1\left(\frac{p_1^r}{1 + \beta_1 \frac{p_1^r}{\alpha p_2^r}} \right), D_2\left(\frac{p_2^r}{1 + \beta_2 \frac{\alpha p_2^r}{p_1^r}} \right) \right). \tag{11}$$

At the solution of (6), the transfer $t^B(p^r)$ equalizes the gains from cooperation across representatives, where gains are weighted by the bargaining power and the tax price of each representative.

For the sake of brevity, we focus on two polar cases: symmetric countries with complete spillovers, and unilateral spillovers. The first case is interesting in that with symmetric countries and representatives, no transfers are made and the level of public good provision is the same as in the basic model without transfers. However, the availability of transfers affect voters' incentives to deviate from a symmetric equilibrium at the electoral stage. Thus, the symmetric case allows us to isolate the impact of transfers on voters' incentives at the electoral stage.

Proposition 9. *Assume $p_1^m = p_2^m$, $D_1 = D_2$, $\alpha = 1$, and $\beta_1 = \beta_2 = 1$. Then there exists $\bar{\varepsilon} \in (1, 2)$ such that if D is more (less) convex than $\bar{D}(p) \equiv p^{-\bar{\varepsilon}}$, any symmetric CEE yields strictly more (less) public good than the NEE. Numerically, $\bar{\varepsilon} \simeq 1.37$.*

Proposition 9 shows that the availability of transfers does not change the qualitative nature of the results of the basic model. First, cooperation can be harmful even in the case of complete spillovers, in which the potential gains from cooperation are greatest. Second, cooperation is beneficial only when the demand function for the public good is sufficiently convex.

The comparison of Propositions 3 and 9 further reveals that cooperation is more likely to be beneficial without transfers. In particular, when D is more convex than p^{-1} and less convex than $p^{-\bar{\varepsilon}}$ for instance when $D(p) = p^{-\varepsilon}$ for some $\varepsilon \in (1, \bar{\varepsilon})$ cooperation is beneficial without transfers but detrimental with transfers. Hence, once voters' incentives are taken into account, the availability of transfers may decrease the gains from cooperation. To see why, consider the incentives of the median voter of country c to marginally increase p_c^r from a symmetric profile $p_c^r = p_{-c}^r$. When transfers are allowed, from (11), this deviation decreases the provision of both public goods by the same amount, but representative c is compensated by a transfer from country $-c$ for his greater cost of provision. When transfers are ruled out, increasing $p_c{}^r$ decreases public good provision to a greater extent, because the greater cost of provision of representative c cannot be compensated by a transfer from country $-c$. Since the median voter of country c cares more about public good consumption than her representative, she is more likely to benefit from the deviation in the former case, because it leads to a smaller decline in public good consumption than in the latter case.

We now turn to the opposite case of a unilateral externality, and assume that $\beta_2 = 0$ while all other parameters can take arbitrary values. If we fix the profile of representatives, transfers are necessary for cooperation to be beneficial in this setup. To see this, note that in the absence of transfer, $\beta_2 = 0$ implies that country 2 has nothing to offer to country 1, so the requirement that cooperation be mutually beneficial implies that for any $p^r \in \mathbb{R}^2_{++}$, $x^B(p^r) = x^N(p^r)$. As a result, the unique CEE is $p^r = p^m$, and it is outcome equivalent to the NEE. When transfers are feasible, one can see from (11) that if we fix the profile of representative p^r, cooperation leaves the level of public good 2 unchanged, but it strictly increases the provision of public good 1, and country 2 makes a transfer to country 1 to compensate its voters for the greater cost of provision. However, the next proposition shows that once voters' incentives are taken into account, cooperation can still be detrimental.

Proposition 10. *Assume $\beta_2 = 0$ while all other parameters are arbitrary. Then any CEE yields strictly less public good in country 2 than the NEE, and if D_1 is more (less) convex than $\overline{D}(p) \equiv p^{-2}$, any CEE yields strictly more (less) public good in country 1.*

Proposition 10 shows that whether cooperation increases the provision of the public good with the spillover effect depends solely on the degree of convexity of its demand function, as in the basic model. However, cooperation always decreases the provision of the public good without the spillover effect. Equivalently, the median voter of country 2 always elects a representative with a higher tax price than herself. This electoral strategy is profitable to her because it decreases the provision of both public goods, but it also decreases the transfer paid by her country to compensate country 1. Since this lower transfer is negotiated by a representative who cares more than her about after tax income relative to public good consumption, it more than offsets the effect of lower public good consumption for her.

The comparison of Propositions 3 and 10 also reveals that when D_1 is more convex than p^{-1} but less convex than p^{-2}, cooperation decreases the provision of both public goods when transfers are available, whereas it has no effect when transfers are ruled out. Thus, as in the symmetric setup, the availability of transfers can make cooperation detrimental.

7. A model with crowding out

As argued in Section 3.1, the main simplifying assumption of the basic model is that voters have separable preferences over the two public goods. This assumption implies that there is no crowding out in the policy making stage of the noncooperative regime: the willingness of a representative to pay for his public good does not depend on the level of public good in the other country. As a result, there is no strategic delegation in the NEE (see Proposition 1). To investigate whether our main results depend on that assumption, we modify the model of Section 3 and consider another widely used public good model in which the contribution of one country crowds out the contribution of the other country.

The policy x_c of each country c is the level of provision of an input (e.g., pollution abatement) that is necessary to produce a public good (e.g., air quality). Voters consume only the public good of their own country, but the quantity g_c they can consume depends both on their own contribution x_c and on the contribution of the other country x_{-c}. Specifically, $g_c = x_c + \beta_{-c}x_{-c}$, and the preferences of voter (i, c) are given by

$$U_{i,c}(x) = t_{i,c}G_c(x_c + \beta_{-c}x_{-c}) - s_{i,c}p_c x_c, \tag{12}$$

where the parameters $t_{i,c}$, $s_{i,c}$, and p_c have the same interpretation as in the basic model. We further assume that $\beta_1\beta_2 \leq 1$.[16] As in the basic model, if we denote $p_{i,c} \equiv s_{i,c}p_c/t_{i,c}$, the preferences of voter (i, c) can equivalently be represented by the following utility function

$$U_c(p_{i,c}, x) = G_c(x_c + \beta_{-c}x_{-c}) - p_{i,c}x_c. \tag{13}$$

If country c was the only contributor (i.e., $x_{-c} = 0$) and voter (i, c) could decide the contribution of her country, she would choose a level of public good $g_c = x_c = (G'c)^{-1}(p_{i,c})$. Thus $D_c \equiv (G'c)^{-1}$ and $p_{i,c}$ can still be interpreted as the public good demand function, and $p_{i,c}$ as the tax price of voter (i, c).

The noncooperative and cooperative electoral equilibria are defined as in the basic model. Since Lemma 1 also holds in this model, the NEE and the CEE are still the Nash equilibria of a game in which the median voters choose the type of their own representative, and the outcome is $x^N(p^r)$ and $x^B(p^r)$, respectively, where x^N and x^B are constructed as in the basic model (see Sections 4.1 and 4.2).

7.1. The noncooperative regime

Let p^r be a profile of representative appointed in the electoral stage of the noncooperative regime. In the policy making stage, if representative p_c^r expects the other country to set a policy x_{-c}, his best response is to set $x_c = \max\{0, D_c(p_c^r) - \beta_{-c}x_{-c}\}$. Thus, in contrast to the basic model, policies are strategic substitutes.

[16] This assumption guarantees that the equilibrium in the policy-making stage of the noncooperative scenario is unique.

Proposition 11. *For any p^r, the unique equilibrium of the policy making stage is*

$$x^N(p^r) = \begin{cases} (0, D_2(p_2^r)) & \text{if } \dfrac{D_1(p_1^r)}{D_2(p_2^r)} \in [0, \beta_2], \\ \left(\dfrac{D_1(p_1^r) - \beta_2 D_2(p_2^r)}{1 - \beta_1\beta_2}, \dfrac{D_2(p_2^r) - \beta_1 D_1(p_1^r)}{1 - \beta_1\beta_2}\right) & \text{if } \dfrac{D_1(p_1^r)}{D_2(p_2^r)} \in \left(\beta_2, \dfrac{1}{\beta_1}\right), \\ (D_1(p_1^r), 0) & \text{if } \dfrac{D_1(p_1^r)}{D_2(p_2^r)} \in \left[\dfrac{1}{\beta_1}, +\infty\right). \end{cases} \tag{14}$$

The equilibrium level of public good in country c is $g_c^N(p^r) = D_c(p_c^r)$ when $x_c^N(p^r) > 0$, and it is $g_c^N(p^r) = \beta_{-c} D_{-c}(p_{-c}^r) > D_c(p_c^r)$ when $x_c^N(p^r) = 0$.

Hence, either $\frac{D_1(p_1^r)}{D_2(p_2^r)} \in \left(\beta_2, \frac{1}{\beta_1}\right)$ that is, representatives have sufficiently homogeneous tax prices and/or β_1 and β_2 are sufficiently small in which case the contributions of both countries are positive, but they partially crowd out each other, or $\frac{D_1(p_1^r)}{D_2(p_2^r)} \notin \left(\beta_2, \frac{1}{\beta_1}\right)$ that is, representatives have sufficiently heterogeneous tax prices and/or β_1 and β_2 are sufficiently large in which case one country free rides and the representative of the other country implements his most preferred contribution.

That the nonnegativity constraints $x_1 \geq 0$ and $x_2 \geq 0$ can bind in the noncooperative policy equilibrium complicates the analysis. The main point of this section is to investigate how strategic substitutability affect the results of the basic model, and from (14), this effect is present at the margin only in interior equilibria. Therefore, for the sake of brevity, the analysis below focuses on electoral equilibria p^r such that $\frac{D_1(p_1^r)}{D_2(p_2^r)} \in \left(\beta_2, \frac{1}{\beta_1}\right)$. This is the case for all NEE and CEE when β_1 and β_2 are sufficiently small. In the working paper version (Loeper, 2015), we consider the opposite case of a pure public good that is, $\beta_1\beta_2 = 1$. In that case, $x^N(p^r)$ is always a corner equilibrium, but the main results are qualitatively unchanged. For convenience, we report these results in Section 7.3.

Proposition 12. *For all p^r such that $\frac{D_1(p_1^r)}{D_2(p_2^r)} \in \left(\beta_2, \frac{1}{\beta_1}\right)$, if p_c^r is a noncooperative electoral best response to p_{-c}^r, then $p_c^r = \frac{p_c^m}{1-\beta_1\beta_2}$, and the corresponding level of public good in country c is $g_c(p^r) = D_c\left(\frac{p_c^m}{1-\beta_1\beta_2}\right)$. Therefore, the unique interior NEE is given by $p^r = \frac{p^m}{1-\beta_1\beta_2}$. In this NEE, both public goods are underprovided: there exists a policy vector x that is strictly preferred to x^{NEE} by a majority of voters in both countries, and any such policy vector yields a higher level of public goods in both countries.*

The intuition for Proposition 12 is as follows. Proposition 11 implies that, in contrast to the basic model, the type p_c^r elected by the voters of country c affects not only their own contribution x_c but also the contribution of the other country x_{-c}. This electoral externality is due to the crowding out effect in the policy making stage. It induces voters to delegate policy making to a higher tax price representative. To understand these electoral incentives, note that Proposition 11 implies that for any interior policy equilibrium $x^N(p^r)$, the level of public good in country c is $g_c = D_c(p_c^r)$, so it depends only on the representative of country c. Thus, in line with the basic model, the electoral stage of the noncooperative regime is strategically equivalent to a game in which the voters of each country c choose g_c, taking g_{-c} as given. However, in contrast to the basic model, because countries' contributions crowd each other out, an extra unit of public good consumption g_c increases that country's contribution x_c by more than one unit. Precisely, (14) implies that the marginal rate of transformation between x_c and g_c is $1 - \beta_1\beta_2$. The voters react to this crowding out effect by inflating the tax price of their representative by a factor $1/(1 - \beta_1\beta_2)$.

7.2. The cooperative regime

With the specification (13), the bargaining program (6) has a unique solution $x^B(p^r)$, and the corresponding level of public good is such that

$$g_c^B(p^r) = x_c^B(p^r) + \beta_{-c} x_{-c}^B(p^r) = D_c\left(\frac{1 - \beta_c B_{-c}/B_c}{1 - \beta_1\beta_2} p_c^r\right), \tag{15}$$

where B_c and B_{-c} are defined as in Section 4.2. By comparing (14) to (15), we see that cooperation requires policy makers to behave as if their tax price was $\frac{1-\beta_c B_{-c}/B_c}{1-\beta_1\beta_2} p_c^r$ instead of p_c^r. In the Appendix A (see Lemma 7), we show that if we fix the representatives' tax price, cooperation unambiguously increases the provision of both public goods, relative to the noncooperative policy equilibrium.

Note that the ratio B_{-c}/B_c plays the same role as in the basic model. If we take it as an exogenous parameter in the formula (15), then as B_{-c}/B_c increases, the policy vector x^B remains Pareto optimal for the two representatives, but x_c^B increases whereas x_{-c}^B decreases, so the outcome of cooperation becomes less favorable to p_c^r and more favorable to p_{-c}^r.

We now turn to the analysis of the electoral stage. The following proposition shows that, as in the basic model (see Proposition 2), once voters' incentives are taken into account, whether cooperation increases public good provision depends on whether the bargaining process treats higher tax price representatives more or less favorably.

Proposition 13. *For all profile of representative* p^r *such that* $\frac{D_1(p_1^r)}{D_2(p_2^r)} \in \left(\beta_2, \frac{1}{\beta_1}\right)$, *if* $p_c{}^r$ *is a cooperative electoral best response to* p_{-c}^r, *then the corresponding level of public good in country c is strictly smaller (greater) than the level of public good in country c in the noncooperative best response of country c to any* $\tilde{p}_{-c}^r$ *if* $\frac{\partial[B_{-c}/B_c]}{\partial p_c^r} < 0\ (>0)$.

The intuition for that result is similar to the intuition for Proposition 2 in the basic model. In the cooperative regime, when choosing the tax price $p_c{}^r$ of their representative, the voters of country c indirectly choose their level of public good consumption g_c: they increase g_c as they lower $p_c{}^r$, as in the noncooperative regime (see Lemma 8 in the Appendix A). However, in contrast to the noncooperative regime, $p_c{}^r$ also affects the level of public good in the other country g_{-c}. For instance, when $\partial[B_{-c}/B_c]/\partial p_c^r < 0$, we see from (15) that $\partial g_{-c}^B/\partial p_c^r > 0$. Thus, as the median voter of country c increases her level of public good consumption $g_c{}^B$ (by decreasing $p_c{}^r$), she decreases g_{-c}^B. This means that the crowding out effect of $x_c{}^B$ on x_{-c}^B is greater in the cooperative regime than in the noncooperative regime: an extra unit of public good consumption in country c requires a greater increases in its contribution x_c. This effect of cooperation induces voters to increase public good provision relative to the NEE.

7.3. When does cooperation increase public good provision?

Proposition 14. *Suppose that for some* $c \in \{1, 2\}$, $D_c(p)$ *is more (less) convex than* $D(p) \equiv \ln(1/p)$ *in the sense of Definition 1.*[17] *Then irrespective of the other parameters of the model, any interior CEE yields more (less) public good in country c than the interior NEE.*

The main finding in Proposition 14 is that as in the basic model, whether cooperation increases public good provision depends only on the convexity of the public good demand function. The only difference is that cooperation increases public good provision in more cases in this model than in the basic model. To see this, note that any function that is more convex than $1/p$ is also more convex than $\ln(1/p)$.

The intuition behind this quantitative difference between the two models is that in this model, the inefficiency of the NEE is worsened by strategic delegation,[18] so it is easier to improve on the NEE. This intuition can be confirmed by considering the pure public good case, that is, $\beta_1\beta_2 = 1$. This case is analyzed formally in the working paper version (Loeper, 2015). We show that in the noncooperative regime, one country always free rides whereas the other country elects its median voter, so strategic delegation does not exacerbates the coordination failure in the NEE. In that pure public good model, cooperation increases (decreases) public good provision when $D_c(p)$ is more (less) convex than $D(p) \equiv 1/p^2$,[19] irrespective of the other parameters of the model. Thus, the main result of the basic model is still qualitatively unchanged, but cooperation is now less likely to increase public good provision relative to the basic model, because any function that is more convex than $1/p^2$ is also more convex than $1/p$.

8. Concluding remarks

It is instructive to relate our results to the empirical literature on the demand for public goods. Most empirical studies have found a relatively small tax price elasticity of the public good demand, typically lower than one (see, e.g., Wildasin, 1987) but estimates greater than one have been found for some particular public goods (see DelRossi and Inman (1999)). In light of Corollary 1, these studies suggest that strategic voting can significantly erode the gains from interjurisdictional Coasian cooperation, and that the impact of cooperation can vary substantially across types of public goods.

It should be noted that most of these estimates are for local public goods. For public goods of national of international scope, existing estimates are based on individual donations. They are not equivalent to tax price elasticities, because charitable donations reflect a desire for warm glow which is crowded out when individual contributions take the form of mandatory taxation. Nevertheless, it is interesting to note that, as for the case of local public goods, these estimates are typically around or below one (see, e.g., Auten et al., 2002, and the references therein), and more importantly, they vary greatly between public goods (Feldstein, 1975; Brooks, 2007).

When empirical estimates of tax price elasticities are not available, our results suggest the following rule of thumb: the negative effect of strategic voting on international cooperation is more severe for public goods whose marginal return decreases more rapidly. Policy expertise can then be used to assess this rate of decrease for a given public good, say fundamental research, disease eradication, or global warming.

Acknowledgements

This paper previously circulated under the title "Public Good Provision and Cooperation Among Representative Democracies." Support from the Ministerio Economía y Competitividad (Spain), Grants ECO 2013 42710 P, MDM 2014 0431, and Comunidad de Madrid, MadEco CM (S2015/HUM 3444) is gratefully acknowledged. Part of this research was conducted while I was research

[17] The demand function $D(p) = \ln(1/p)$ corresponds to the public good technology $G(x) = -\exp(-x)$, as assumed in Segendorff (1998).

[18] Public goods are underprovided in the NEE not only because elected representatives do not internalize externalities, but also because they are less willing to provide them than the median voters that elected them. Formally, there exists $x \in \mathbb{R}_+^2$ such that x is majority preferred in both countries to $x^N(p^m)$, and any such policy vector x is such that $x_1 > x_1^N(p^m)$ and $x_2 > x_2^N(p^m)$. And $x^N(p^m)$ is in turn majority preferred in both countries to $x^{NEE}\left(\frac{p^m}{1-\beta_1\beta_2}\right)$.

[19] The demand function $D(p) = 1/p^2$ corresponds to the public good technology $G(x) = \sqrt{x}$, as assumed in Gradstein (2004).

fellow at Banco de Espana. I thank three anonymous referees for their very valuable comments. All remaining errors are mine.

Appendix A

Proof of Lemma 1 For each $c \in \{1, 2\}$, the utility functions in (2) and in (13) are linear in $p_{i,c}$. Therefore, they satisfy the intermediate preference property of Grandmont (1978). That is, for any voters i, j, k in country c, if j's type is in between that of i and k i.e., if $p_{i,c} \leq p_{j,c} \leq p_{k,c}$ or $p_{i,c} \geq p_{j,c} \geq p_{k,c}$ then when i and k agree on how to rank two given policies $x, y \in \mathbb{R}_+^2$, j must also agree with i and k. This property implies that the majority preferences in country c coincide with the preferences of p_c^m (Grandmont, 1978).

A.1. Proofs for Section 4.1

Under our assumptions on G_c, $U_c(p_c^r, x)$ is strictly concave and differentiable in $\boldsymbol{x_c}$, strictly decreasing as $x_c \to 0$, and strictly increasing as $x_c \to \infty$. Therefore, its unique maximum $\boldsymbol{x_c}^{\mathcal{N}}$ is characterized by the F.O.C. $\partial U_c / \partial x_c = 0$, which implies $x_c^N(p^r) = D_c(p_c^r)$ and proves (4).

Proof of Proposition 1. From Lemma 1, the noncooperative electoral best response of the voters of country c to some p_{-c}^r is the representative's type p_c^r most preferred by the median voter of country c, given the policy mapping $p_c^r \to x^N(p_c^r, p_{-c}^r)$. So it is the solution to $\max_{p_c^r > 0} U_c(p_c^m, x^N(p^r))$. Substituting $x^N(p^r) = (D_1(p_1^r), D_2(p_2^r))$ in this program, we see that the objective function is strictly quasi concave, and the F.O.C. yields $p_c^r = p_c^m$, irrespective of p_{-c}^r, which proves the first part of Proposition 1.

As argued in Section 4.1, the policy outcome of the NEE x^{NEE} is the outcome of the game in which the median voter of each country c controls x_c, taking x_{-c} as given. In the presence of externalities, it is obvious that the Nash equilibrium of that game is inefficient, so there exists a policy vector x^* that is strictly preferred by both median voters. From Lemma 1, x^* is strictly preferred by a majority of voters in both countries as well. To conclude the proof, it remains to show that any such vector x^* must be such that $x_1^* > D_1(p_1^m)$ and $x_2^* > D_2(p_2^m)$. Suppose that, say, the latter inequality is violated. Then

$$U_1(p_1^m, x^*) < U_1(p_1^m, x_1^*, D_2(p_2^m)) \leq \max_{x_1 \geq 0} U_1(p_1^m, x_1, D_2(p_2^m)) = U_1(p_1^m, D_1(p_1^m), D_2(p_2^m)),$$

so the median voter of country 1 does not strictly prefer x^* to x^{NEE}, a contradiction.□

A.2. Proofs for Section 4.2

Lemma 2. *Properties (9) and (10) imply that the program (6) has a unique solution, and it is interior.*

Proof. We first prove that (10) implies that B is strictly quasi concave. For all $\Delta', \Delta'' \in \mathbb{R}_{++}^2$, and all $\alpha \in [0, 1]$, let $\Phi(\alpha) \equiv B_1(\alpha\Delta' + (1 - \alpha)\Delta'')$. Note first that B is strictly quasi concave if and only if Φ is strictly quasi concave for all $\Delta', \Delta'' \in \mathbb{R}_{++}^2$ with $\Delta' \neq \Delta''$. Moreover,

$$\frac{\partial \Phi}{\partial \alpha} = (\Delta'1 - \Delta''1)\frac{\partial B}{\partial \Delta_1}(\alpha\Delta' + (1 - \alpha)\Delta'') + (\Delta'2 - \Delta''2)\frac{\partial B}{\partial \Delta_2}(\alpha\Delta' + (1 - \alpha)\Delta'') = \frac{\partial B}{\partial \Delta_1}(\alpha\Delta' + (1 - \alpha)\Delta'') \left[(\Delta'1 - \Delta''1) + (\Delta'2 - \Delta''2)\frac{\frac{\partial B}{\partial \Delta_2}(\alpha\Delta' + (1 - \alpha)\Delta'')}{\frac{\partial B}{\partial \Delta_1}(\alpha\Delta' + (1 - \alpha)\Delta'')} \right].$$

If $\Delta'1 \geq \Delta''1$ and $\Delta'2 \geq \Delta''2$ with one inequality strict, then the above equation shows that $\partial\Phi/\partial\alpha$ has the same strict sign for all $\alpha \in [0, 1]$. Therefore, Φ is strictly monotonic, and thus strictly quasi concave. The same conclusion applies if $\Delta'1 \leq \Delta''1$ and $\Delta'2 \leq \Delta''2$ with one inequality strict. If $\Delta'1 > \Delta''1$ and $\Delta'2 < \Delta''2$, then from (10), $\frac{\frac{\partial B}{\partial \Delta_2}(\alpha\Delta' + (1 - \alpha)\Delta'')}{\frac{\partial B}{\partial \Delta_1}(\alpha\Delta' + (1 - \alpha)\Delta'')}$ is strictly increasing in α, so the term in bracket on the R.H.S. of the above expression is strictly decreasing in α, so $\frac{\partial \Phi}{\partial \alpha}$ satisfies the strict single crossing condition in α, and Φ is strictly quasi concave. An analogous reasoning leads to the same conclusion in the remaining case $\Delta'1 < \Delta''1$ and $\Delta'2 > \Delta''2$.

We now prove that (6) has a unique solution. Note that $U_c(p_c, x)$ and therefore $\Delta_c(p^r, x)$ are strictly concave in x. Since B is strictly quasi concave, $x \to B(\Delta_1(p^r, x)), \Delta_2(p^r, x))$ is strictly quasi concave, so (6) has at most one solution. As $x_1 \to \infty$ or as $x_2 \to \infty$, either $\Delta_1(p^r, x) \to -\infty$ or $\Delta_2(p^r, x) \to -\infty$, so Property (9) implies that w.l.o.g., we can restrict the choice set of the program (6) to a bounded and closed subset of $\mathbb{R}_+^2$, which proves that (6) has at least one solution.

Since $\lim_{x \to 0} G'c(x)$=+∞, one can easily see that any Pareto optimal $x \in \mathbb{R}_+^2$ must be such that $x_1 > 0$ and $x_2 > 0$. Since B is strictly increasing in each coordinate, $x^B(p^r)$ must be Pareto optimal, and thus interior.

From (4), the welfare gains for representative $\boldsymbol{p_c}^{\boldsymbol{r}}$ from implementing policy x instead of the noncooperative equilibrium $x^N(p^r)$ is

$$\Delta_c(p^r, x) = G_c(x_c) + \beta_{-c}G_{-c}(x_{-c}) - p_c^r x_c - G_c(D_c(p_c^r)) - \beta_{-c}G_{-c}(D_{-c}(p_{-c}^r)) + p^r D_c(p_c^r). \quad (16)$$

The following Lemma establishes inter alia Eq. (7).

Lemma 3. *If we denote* $\hat{x}(f,p^r) \equiv \left(D_1\left(\frac{p_1^r}{1+\beta_1 f}\right), D_2\left(\frac{p_2^r}{1+\beta_2/f}\right)\right)$, *then the equation*

$$f = \frac{\frac{\partial B}{\partial \Delta_2}(\Delta(p^r, \hat{x}(f,p^r)))}{\frac{\partial B}{\partial \Delta_1}(\Delta(p^r, \hat{x}(f,p^r)))} \tag{17}$$

has a unique solution f^*, *and* $x^B(p^r) = \hat{x}(f^*, p^r)$.

Proof. From Lemma 2, $x^B(p^r)$ is the unique interior solution to the program (6), which is characterized by the F.O.C. $\frac{\partial[B(\Delta(p^r,x))]}{\partial x} = 0$. Using (16), this F.O.C. becomes

$$\frac{\partial B}{\partial \Delta_c}(\Delta(p^r,x))(G'c(x_c) - p_c^r) + \frac{\partial B}{\partial \Delta_{-c}}(\Delta(p^r,x))\beta_c G'c(x_c) = 0. \tag{18}$$

Isolating $G'c(x_c)$ in (18) and using the notation B_c introduced in (8), we obtain $G'c(x_c) = \frac{p_c^r}{1+\beta_c B_{-c}/B_c}$. Applying the function D_c on both sides of the latter equation, we obtain (7), which can be rewritten as $x^B(p^r) = \hat{x}(f,p^r)$, where $f = \frac{B_{-c}(p^r)}{B_c(p^r)} = \frac{\frac{\partial B}{\partial \Delta_2}(\Delta(p^r,\hat{x}(f,p^r)))}{\frac{\partial B}{\partial \Delta_1}(\Delta(p^r,\hat{x}(f,p^r)))}$, so f is a solution to (17).

Reciprocally, let f^* be a solution to (17), and let $x^* = \hat{x}(f^*, p^r)$. Then using successively the definition of $\hat{x}$ and (17), we obtain

$$\begin{aligned}\frac{\partial B}{\partial \Delta_1}(\Delta(p^r,x^*))(G'1(x_1^*) - p_1^r) + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r,x^*))\beta_1 G'1(x_1^*) &= \frac{\partial B}{\partial \Delta_1}(\Delta(p^r,x^*))\left(\frac{p_1^r}{1+\beta_1 f^*} - p_1^r\right) + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r,x^*))\beta_1 \frac{p_1^r}{1+\beta_1 f^*} \\ &= -\frac{\partial B}{\partial \Delta_1}(\Delta(p^r,x^*))f^* + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r,x^*)) = 0,\end{aligned}$$

which shows that x^* satisfies the F.O.C. (18) for c=1. An analogous reasoning shows that it also satisfies (18) for $c = 2$, so x^* is a solution of (6), and $x^* = x^B(p^r)$. Thus, we have shown that the solutions to (17) are in a one to one relationship with the solutions of the program (6). Since the latter program has a unique solution, so does (17).□

The next lemma shows that in the cooperative regime, the contribution x_c^B of each country c is strictly decreasing in the type of its representative p_c^r, and thus that as claimed in Section 4.2, by choosing p_c^r, the voters of each country effectively chooses their contribution x_c^B (although p_c^r also affects the contribution of the other country x_{-c}^B, see Lemma 5 below).

Lemma 4. $\frac{\partial x_c^B}{p_c^r} < 0$.

Proof. Suppose by contradiction that there exists $\hat{p}^r$ such that $\frac{\partial x_c^B}{p_c^r}(\hat{p}^r) \geq 0$. From (7),

$$\frac{\partial x_c^B}{\partial p_c^r} = D'c\left(\frac{p_c^r}{1+\beta_c B_{-c}/B_c}\right) \times \left(\frac{1}{1+\beta_c B_{-c}/B_c} - \frac{\beta_c p_c^r}{(1+\beta_c B_{-c}/B_c)^2}\frac{\partial[B_{-c}/B_c]}{\partial p_c^r}\right)$$

Since $D_c < 0$ and $\frac{\partial x_c^B}{p_c^r}(\hat{p}^r) \geq 0$, the above equation implies that

$$\frac{\partial[B_{-c}/B_c]}{\partial p_c^r}(\hat{p}^r) > 0. \tag{19}$$

Together with (7), (19) implies that $\frac{\partial x_{-c}^B}{\partial p_c^r}(\hat{p}^r) < 0$. Using (7) again and $D_c = (G'c)^{-1}$, we obtain that for all p_c^r,

$$\frac{\partial[U_c(\hat{p}_c^r, x^B(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r} = \frac{p_c^r}{1+\beta_c B_{-c}/B_c}\frac{\partial x_c^B}{\partial p_c^r} + \frac{\beta_{-c}\hat{p}_{-c}^r}{1+\beta_{-c}B_c/B_{-c}}\frac{\partial x_{-c}^B}{\partial p_c^r} - \hat{p}_c^r\frac{\partial x_c^B}{\partial p_c^r} = \frac{-p_c^r\beta_c B_{-c}/B_c}{1+\beta_c B_{-c}/B_c}\frac{\partial x_c^B}{\partial p_c^r} + \frac{\beta_{-c}\hat{p}_{-c}^r}{1+\beta_{-c}B_c/B_{-c}}\frac{\partial x_{-c}^B}{\partial p_c^r} + (p_c^r - \hat{p}_c^r)\frac{\partial x_c^B}{\partial p_c^r}$$

where the functions in the above expressions are evaluated at $(p_c^r, \hat{p}_{-c}^r)$. If we set $p_c^r = \hat{p}_c^r$ and substitute $\frac{\partial x_c^B}{p_c^r}(\hat{p}^r) \geq 0$ and $\frac{\partial x_{-c}^B}{\partial p_c^r}(\hat{p}^r) < 0$ into the above expression, we obtain that $\frac{\partial[U_c(\hat{p}_c^r, x^B(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r}(\hat{p}^r) < 0$. Since $x^B(\hat{p}^r)$ is Pareto optimal for $\hat{p}^r$, this implies that $\frac{\partial[U_{-c}(\hat{p}_c^r, x^B(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r}(\hat{p}^r) \geq 0$. Using (4), we always have $\frac{\partial[U_{-c}(\hat{p}_{-c}^r, x^N(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r}(\hat{p}^r) < 0$. Using notation (5), the latter two inequalities imply that the gain from cooperation increases for $\hat{p}_{-c}^r$ as p_c^r increases around $\hat{p}_c^r$, that is,

$$\frac{\partial[\Delta_{-c}(p^r, x^N(p^r))]}{\partial p_c^r}(\hat{p}^r) > 0. \tag{20}$$

Differentiating (16) and using $D_c = (G'c)^{-1}$, we obtain

$$\frac{\partial[\Delta_c(p^r, x^N(p^r))]}{\partial p_c^r} = \frac{p_c^r}{1+\beta_c B_{-c}/B_c}\frac{\partial x_c^B}{\partial p_c^r} + \frac{\beta_{-c} p_{-c}^r}{1+\beta_{-c} B_c/B_{-c}}\frac{\partial x_{-c}^B}{\partial p_c^r} - p_c^r \frac{\partial x_c^B}{\partial p_c^r} - x_c^B + D_c(p_c^r) - \frac{p_c^r \beta_c B_{-c}/B_c}{1+\beta_c B_{-c}/B_c}\frac{\partial x_c^B}{\partial p_c^r} + \frac{\beta_{-c} p_{-c}^r}{1+\beta_{-c} B_c/B_{-c}}\frac{\partial x_{-c}^B}{\partial p_c^r} + D_c(p_c^r) - x_c^B. \tag{21}$$

From (4), $D_c(p_c^r) < x_c^B$. Substituting the latter inequality, $\frac{\partial x_c^B}{p_c^r}(\widehat{p}^r) \geq 0$, and $\frac{\partial x_{-c}^B}{\partial p_c^r}(\widehat{p}^r) < 0$ into (21), we obtain that $\frac{\partial[\Delta_c(p^r, x^N(p^r))]}{\partial p_c^r}(\widehat{p}^r) < 0$. Using property (10), the latter inequality and (20) imply then that B_{-c}/B_c must be strictly decreasing in p_c^r at $p^r = \widehat{p}^r$, a contradiction with (19).□

Proof of Proposition 2. *Step 1*: *As* $p_c^r \to +\infty$, $x_c^B(p^r) \to 0$ *and* $x_{-c}^B(p^r) \to D(p_{-c}^r)$.Suppose by contradiction that $\lim_{n\to+\infty} x_c^B(p_c^r(n), p_{-c}^r) > 0$ for some sequence $p_c^r(n) \to +\infty$. From (7), this implies that as $n \to \infty$, $\frac{B_{-c}(p_c^r(n), p_{-c}^r)}{B_c(p_c^{r,n}, p_{-c}^r)} \to +\infty$, and therefore, $x_{-c}^B(p_c^r(n), p_{-c}^r) \to D_{-c}(p_{-c}^r)$. This implies in turn that $\Delta_c((p_c^r(n), p_{-c}^r), x^B(p_c^r(n), p_{-c}^r)) \to -\infty$, which contradicts property (9), and thus proves the first limit of Step 1. Since $\lim_{p_c^r\to+\infty} x_c^N(p^r) = 0$, the second limit of Step 1 follows then directly from the first limit together with (9) for country c.

Step 2: *As* $p_c^r \to 0$, $x_c^B(p^r) \to +\infty$, $x_{-c}^B(p^r) = o(x_c^B(p^r))$, *and* $U_c(p_c^m, x_c^B(p^r)) \to -\infty$.Using (7) and our assumptions on G_c, $x_c^B(p^r) \geq D_c(p_c^r) \to_{p_c^r\to 0} +\infty$, which proves the first limit of Step 2. To prove the second limit of Step 2, suppose by contradiction that $x_{-c}^B(p^r)$ is not $o(x_c^B(p^r))$ as $p_c^r \to 0$. Then there exists a sequence $p_c^n \to 0$ such that $x_{-c}^B(p_c^n, p_{-c}^r) \geq A x_c^B(p_c^n, p_{-c}^r)$ for some $A > 0$. Therefore, $x_{-c}^B(p_c^n, p_{-c}^r) \to +\infty$. Since $\lim_{x\to+\infty} G'c(x) = 0$, this implies that as $p_c^r \to 0$, $G'c(x_c^B(p_c^n, p_{-c}^r)) = o(x_{-c}^B(p_c^n, p_{-c}^r))$ and $G' - c(x_{-c}^B(p_c^n, p_{-c}^r)) = o(x_{-c}^B(p^r))$, so from (2), $U_{-c}(p_{-c}^r, x^B(p_c^n, p_{-c}^r)) \to -\infty$, and therefore, $\Delta_{-c}(p^r, x^B(p_c^n, p_{-c}^r)) \to -\infty$, which contradicts property (9).

To prove the third limit, note that since $\lim_{x_c\to+\infty} G'c(x_c) = 0$, the first limit of Step 2 implies that as $p_c^r \to 0$, $G_c(x_c^B(p^r)) = o(x_c^B(p^r))$. Since $\lim_{x_{-c}\to+\infty} G' - c(x_{-c}) = 0$, the second limit implies that as $p_c^r \to 0$, $G_{-c}(x_{-c}^B(p^r)) = o(x_{-c}^B(p^r)) = o(x_c^B(p^r))$. Therefore, $\lim_{p_c\to 0} U_c(p_c^m, x_c^B(p^r)) = -\infty$.

Step 3: *a cooperative electoral best response exists.*

From Definition 3 and Lemma 1, the cooperative electoral best response of the voters of country c to some p_{-c}^r are the solutions of

$$\max_{p_c^r > 0} U_c(p_c^m, x_c^B(p^r)). \tag{22}$$

The last limit in Step 2 implies that w.l.o.g., we can restrict the choice set of the program (22) to $[\underline{b}, +\infty)$ for some $\underline{b} > 0$. From Step 1, $\lim_{p_c\to+\infty} x^B(p^r) = (0, D_{-c}(p_{-c}^r))$, and from (7), for all finite p_c, $x_c^B(p^r) > 0$ and $x_{-c}^B(p^r) > D_{-c}(p_{-c}^r)$. Note that $\frac{\partial U_c}{\partial x_{-c}}(p_c^m, x) > 0$ and for x_c sufficiently small, $\frac{\partial U_c}{\partial x_c}(p_c^m, x) > 0$, so w.l.o.g., we can further restrict the choice set of the program (22) to $[\underline{b}, \bar{b}]$ for some $\bar{b} > \underline{b}$. Since $U_1(p_1^m, x^B(p^r))$ is continuous, (22) has a solution.

Step 4: *if* p_c^r *is a cooperative electoral best response to* p_{-c}^r, *and if* $\widetilde{p}_c^r$ *is a noncooperative electoral best response to some* $\widetilde{p}_{-c}^r$, *then* $x_c^B(p^r) - x_c^N(\widetilde{p}^r)$ *has the same sign as* $\partial[B_{-c}/B_c]/\partial p_c^r$.

Any solution to (22) must satisfy the F.O.C. $\partial[U_c(p_c^m, x_c^B(p^r))]/\partial p_c^r = 0$. Using (7) and $D_c = (G'c)^{-1}$, we obtain

$$\frac{\partial[U_c(p_c^m, x_c^B(p^r))]}{\partial p_c^r} = \frac{p_c^r}{1+\beta_c \dfrac{B_{-c}(p^r)}{B_c(p^r)}}\frac{\partial x_c^B}{\partial p_c^r} + \frac{\beta_{-c} p_{-c}^r}{1+\beta_{-c}\frac{B_c(p^r)}{B_{-c}(p^r)}}\frac{\partial x_{-c}^B}{\partial p_c^r} - p_c^m \frac{\partial x_c^B}{\partial p_c^r}. \tag{23}$$

Suppose $\overline{p}_c^r$ is a cooperative electoral best response to p_{-c}^r. Substituting the F.O.C. into the above equation and rearranging terms, we obtain

$$\frac{\overline{p}_c^r}{1+\beta_c \frac{B_{-c}(\overline{p}_c^r, p_{-c}^r)}{B_c(\overline{p}_c^r, p_{-c}^r)}} = p_c^m - \frac{\beta_{-c} p_{-c}^r}{1+\beta_{-c}\dfrac{B_c(\overline{p}_c^r, p_{-c}^r)}{B_{-c}(\overline{p}_c^r, p_{-c}^r)}} \frac{\dfrac{\partial x_{-c}^B}{\partial p_c^r}(\overline{p}_c^r, p_{-c}^r)}{\dfrac{\partial x_c^B}{\partial p_c^r}(\overline{p}_c^r, p_{-c}^r)}, \tag{24}$$

where the division by $\partial x_c^B/\partial p_c^r$ is valid since from Lemma 4, $\partial x_c^B/\partial p_c^r > 0$. Substituting (24) into (7), we obtain

$$x_c^B(p^r) = D_c\left(p_c^m - \frac{\beta_{-c} p_{-c}^r}{1+\beta_{-c}\dfrac{B_c(p^r)}{B_{-c}(p^r)}} \frac{\dfrac{\partial x_{-c}^B}{\partial p_c^r}(p^r)}{\dfrac{\partial x_c^B}{\partial p_c^r}(p^r)}\right).$$

From Proposition 1, irrespective of $\widetilde{p}_{-c}^r$, $x_c^N(\widetilde{p}^r) = D_c(p_c^m)$. The two latter equations imply that $x_c^B(p^r) - x_c^N(\widetilde{p}^r)$ has the same sign as $\frac{\partial x_{-c}^B}{\partial p_c^r}(p^r)/\frac{\partial x_c^B}{\partial p_c^r}(p^r)$, which from Lemma 4 has the opposite sign of $\frac{\partial x_{-c}^B}{\partial p_c^r}(p^r)$, which from (7), has the opposite sign of $\frac{\partial[B_{-c}/B_c]}{\partial p_c^r}$, as needed.

A.3. Proofs for Section 4.3

The next lemma shows that the curvature of D_c determines how the representative's type affects B_{-c}/B_c.

Lemma 5. *Let ϕ_c be such that for all $p_c^r > 0$, $D_c(p_c^r) = \phi_c(1/p_c^r)$. If ϕ_c is strictly convex (concave), then $\partial[B_{-c}/B_c]/\partial p_c^r > 0$ (<0).*

Proof. We prove the Lemma for $c = 1$, the proof for c=2 is analogous. From Lemma 3, B_2/B_1 is given by the solution f to the fixed point condition (17) for c=1. Therefore, it suffices to show that if ϕ_1 is strictly convex (concave), then $\partial f/\partial p_1^r > 0$ (<0).

Differentiating (17) w.r.t. p_1^r, , we obtain

$$\frac{\partial f}{\partial p_1^r} = \frac{\dfrac{\partial\left[\dfrac{\partial B}{\partial \Delta_2}(\Delta(p^r, \hat{x}(f, p^r)))\right]}{\partial p_1^r}\dfrac{\partial B}{\partial \Delta_1}(\Delta(p^r, \hat{x}(f, p^r))) - \dfrac{\partial B}{\partial \Delta_2}(\Delta(p^r, \hat{x}(f, p^r)))\dfrac{\partial\left[\dfrac{\partial B}{\partial \Delta_1}(\Delta(p^r, \hat{x}(f, p^r)))\right]}{\partial p_1^r}}{\left(\dfrac{\partial B}{\partial \Delta_1}(\Delta(p^r, \hat{x}(f, p^r)))\right)^2}.$$

By definition of $\hat{x}$ (see Lemma 3) $\partial\hat{x}_2/\partial p_1^r = 0$, so the above equation can be expanded as follows

$$\begin{aligned}\left(\frac{\partial B}{\partial \Delta_1}\right)^2\frac{\partial f}{\partial p_1^r} &= \frac{\partial B}{\partial \Delta_1}\frac{\partial^2 B}{\partial \Delta_1\partial \Delta_2}\left(\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f}\frac{\partial f}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}\frac{\partial f}{\partial p_1^r}\right)\\
&\quad + \frac{\partial B}{\partial \Delta_1}\frac{\partial^2 B}{(\partial \Delta_2)^2}\left(\frac{\partial \Delta_2}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f}\frac{\partial f}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}\frac{\partial f}{\partial p_1^r}\right)\\
&\quad - \frac{\partial B}{\partial \Delta_2}\frac{\partial^2 B}{(\partial \Delta_1)^2}\left(\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f}\frac{\partial f}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}\frac{\partial f}{\partial p_1^r}\right)\\
&\quad - \frac{\partial B}{\partial \Delta_2}\frac{\partial^2 B}{\partial \Delta_1\partial \Delta_2}\left(\frac{\partial \Delta_2}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f}\frac{\partial f}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}\frac{\partial f}{\partial p_1^r}\right),\end{aligned}$$

where the derivatives of B in the above equation are all evaluated at $\Delta = \Delta(p^r, \hat{x}(f, p^r))$. Solving for $\partial f/\partial p_1^r$, we obtain

$$\frac{\partial f}{\partial p_1^r} = \frac{\begin{aligned}&\left(\frac{\partial B}{\partial \Delta_1}\frac{\partial^2 B}{\partial \Delta_1\partial \Delta_2} - \frac{\partial B}{\partial \Delta_2}\frac{\partial^2 B}{(\partial \Delta_1)^2}\right)\left(\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}\right)\\ &-\left(\frac{\partial B}{\partial \Delta_2}\frac{\partial^2 B}{\partial \Delta_1\partial \Delta_2} - \frac{\partial B}{\partial \Delta_1}\frac{\partial^2 B}{(\partial \Delta_2)^2}\right)\left(\frac{\partial \Delta_2}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}\right)\end{aligned}}{\begin{aligned}&\left(\frac{\partial B}{\partial \Delta_1}\right)^2 - \left(\frac{\partial B}{\partial \Delta_1}\frac{\partial^2 B}{\partial \Delta_1\partial \Delta_2} - \frac{\partial B}{\partial \Delta_2}\frac{\partial^2 B}{(\partial \Delta_1)^2}\right)\left(\frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f} + \frac{\partial \Delta_1}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}\right)\\ &+\left(\frac{\partial B}{\partial \Delta_2}\frac{\partial^2 B}{\partial \Delta_1\partial \Delta_2} - \frac{\partial B}{\partial \Delta_1}\frac{\partial^2 B}{(\partial \Delta_2)^2}\right)\left(\frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f} + \frac{\partial \Delta_2}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}\right)\end{aligned}}. \tag{25}$$

Since $D'c < 0$, by definition of $\hat{x}$ (see Lemma 3), we have $\frac{\partial \hat{x}_1}{\partial f} > 0$ and $\frac{\partial \hat{x}_2}{\partial f} < 0$. Moreover, $\frac{\partial \Delta_1}{\partial x_1}(p_1^r, \hat{x}_1) = G'(\hat{x}_1) - p_1^r = \frac{p_1^r}{1+\beta_1 f} - p_1^r < 0$, and $\frac{\partial \Delta_1}{\partial x_2}(p_1^r, \hat{x}_1) = G'(\hat{x}_1) > 0$. Therefore, the term $\frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f} + \frac{\partial \Delta_1}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}$ in the denominator of the R.H.S. of (25) is negative. An analogous reasoning implies that $\frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial f} + \frac{\partial \Delta_2}{\partial x_2}\frac{\partial \hat{x}_2}{\partial f}$ is positive. Using property (10), we have

$$\frac{\partial B}{\partial \Delta_2}\frac{\partial^2 B}{\partial \Delta_1\partial \Delta_2} - \frac{\partial B}{\partial \Delta_1}\frac{\partial^2 B}{(\partial \Delta_2)^2} = \frac{\partial\left[\frac{\partial B}{\partial \Delta_1}/\frac{\partial B}{\partial \Delta_2}\right]}{\partial \Delta_2}\left(\frac{\partial B}{\partial \Delta_2}\right)^2 > 0,$$

and the same inequality hold if we reverse the role of 1 and 2. Therefore, the denominator of the R.H.S. of (25) is strictly positive, so $\frac{\partial f}{\partial p_1^r}$ has the same sign as its numerator. Thus, we have shown that $\frac{\partial f}{\partial p_1^r}$ is positive (negative) if $\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}$ and $-\left(\frac{\partial \Delta_2}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}\right)$ are both positive (negative). Therefore, to conclude the proof, it suffices to show that $\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}$ and $-\left(\frac{\partial \Delta_2}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}\right)$ are positive (negative) when ϕ_1 is strictly convex (concave).

By definition of ϕ_1, for all $p_1 > 0$, $\phi_{\prime 1}(1/p_1) = -(p_1)^2 D'1(p_1)$. Using the latter identity, (16) and the definition of $\hat{x}$ (see Lemma 3), we obtain

$$\frac{\partial \Delta_2}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r} = -\beta_1 D'1(p_1^r)p_1^r + \frac{\beta_1 G'1(\hat{x}_1)}{1+\beta_1 f}D'1\left(\frac{p_1^r}{1+\beta_1 f}\right)\frac{\beta_1}{p_1^r}\left(\phi'1\left(\frac{1}{p_1^r}\right) - \phi'1\left(\frac{1+\beta_1 f}{p_1^r}\right)\right), \tag{26}$$

and

$$\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r} = -\hat{x}_1 - p_1^r D'1(p_1^r) + D_1(p_1^r) + p_1^r D'1(p_1^r) + \frac{G'(\hat{x}_1) - p_1^r}{1+\beta_1 f} D'1\left(\frac{p_1^r}{1+\beta_1 f}\right)$$
$$= -\phi_1\left(\frac{1+\beta_1 f}{p_1^r}\right) + \phi_1\left(\frac{1}{p_1^r}\right) + \frac{\beta_1 f}{p_1^r}\phi'1\left(\frac{1+\beta_1 f}{p_1^r}\right). \tag{27}$$

The intermediate value theorem implies that there exists $p \in \left(\frac{p_1^r}{1+\beta_1 f}, p_1^r\right)$ such that $\phi_1\left(\frac{1+\beta_1 f}{p_1^r}\right) - \phi_1\left(\frac{1}{p_1^r}\right) = \frac{\beta_1 f}{p_1^r}\phi'_1\left(\frac{1}{p}\right)$. Substituting the latter identity into (27), we obtain

$$\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r} = \frac{\beta_1 f}{p_1^r}\left[\phi'1\left(\frac{1+\beta_1 f}{p_1^r}\right) - \phi'1\left(\frac{1}{p}\right)\right] \tag{28}$$

Since $\frac{1}{p_1^r} < \frac{1}{p} < \frac{1+\beta_1 f}{p_1^r}$, (26) and (28) are both positive (negative) when ϕ_1 is strictly increasing (decreasing), as needed.[20]

Proposition 3 follows readily from the following proposition, and Definitions 2 and 3.□

Proposition 15. *Let $p^r \in (0, +\infty)^2$ and $\tilde{p}^r \in (0, +\infty)^2$ be two profiles of representatives such that p_c^r is a cooperative electoral best response of the voters of country c to p_{-c}^r, and $\tilde{p}_c^r$ is a noncooperative electoral best response to $\tilde{p}_{-c}^r$. If D_c is more (less) convex than the unit elastic demand, then $x_c^B(p^r)$ is greater (smaller) than $x_c^N(\tilde{p}^r)$.*

Proof. From Proposition 2, a cooperative electoral best response p_c^r to some p_{-c}^r yields more (less) public good in country c than the noncooperative electoral best response $\tilde{p}_c^r$ to some $\tilde{p}_{-c}^r$ if $\frac{\partial[B_{-c}/B_c]}{\partial p_c^r}$ evaluated at p^r is positive (negative). Using Lemma 5, this is the case when ϕ_c is strictly convex (concave). The conclusion follows from Definition 1.

Proof of Corollary 1. In the isoelastic case, $D_c(p) = p^{-\varepsilon_c}$ so $\phi_c(p) = p^{\varepsilon_c}$, and ϕ_c is strictly convex (concave) when $\varepsilon_c > 1$ ($\varepsilon_c < 1$). The corollary follows then readily from Proposition 15.

A.4. Proofs for Section 4.4

Proposition 4 is a direct corollary of the following proposition, and of Definitions 2 and 3.

Proposition 16. *Suppose that for some $c \in \{1, 2\}$, D_c is less convex than the unit elastic demand, and let $p^r \in (0, +\infty)^2$ be such that p_c^r is a cooperative electoral best response of the voters of country c to p_{-c}^r. Then the corresponding outcome $x^B(p^r)$ makes a majority of voters in country c strictly worse off relative to the outcome of the NEE $x^N(p^{NEE})$.*

Proof. Under the condition of the proposition, Proposition 15 implies that $x_c^B(p^r) < x_c^N(p^{NEE})$. To complete the proof, in what follows, we show that for all $y \in \mathbb{R}_+^2$, if $y_c < x_c^N(p^{NEE})$, then a majority of voters in country c strictly prefer $x^N(p^{NEE})$ to y.

Since $y_c < x_c^N(p^{NEE})$, the median voter in country $-c$ strictly prefers $x^N(p^{NEE})$ to $(y_c, x_{-c}^N(p^{NEE}))$. Note that $x_{-c} \to U_{-c}(p_{-c}^m, y_c, x_{-c})$ is maximized at $x_{-c} = D_{-c}(p_{-c}^m) = x_{-c}^N(p^{NEE})$. Therefore, the median voter in country c strictly prefers $(y_c, x_{-c}^N(p^{NEE}))$ to y. By transitivity, she strictly prefers $x^N(p^{NEE})$ to y. The proposition follows then from Lemma 1.□

Proof of Proposition 5. Suppose first that $p_c^{CEE} \leq p_c^m$ for some country c. Since p_c^{CEE} is a best response to p_{-c}^{CEE}, the median voter in country c prefers $x^B(p^{CEE})$ to $x^B(p_c^{CEE}, p_{-c}^m)$, and from condition 9, she also prefers $x^B(p_c^{CEE}, p_{-c}^m)$ to $x^N(p_c^{CEE}, p_{-c}^m) = (D_c(p_c^{CEE}), D_{-c}(p_{-c}^m))$. Since $p_c^{CEE} \leq p_c^m$, we have $D_c(p_c^{CEE}) \geq D_c(p_c^m)$, so she also prefers $(D_c(p_c^{CEE}), D_{-c}(p_{-c}^m))$ to $x^N(p^{NEE}) = (D_c(p_c^m), D_{-c}(p_{-c}^m))$. By transitivity, she strictly prefers $x^B(p^{CEE})$ to $x^N(p^{NEE})$. Proposition 5 follows then from Lemma 1.

Suppose now that $p_1^{CEE} > p_1^m$ and $p_2^{CEE} > p_2^m$. Since D_1 are D_2 are less convex than a unit elastic demand, Proposition 3 implies then that for all $c \in \{1, 2\}$,

$$x_c^B(p^m) > x_c^B(p^{CEE}) > x_c^N(p^m). \tag{29}$$

For all $\alpha \in [0, 1]$, let $x(\alpha) \equiv (1-\alpha)x_c^B(p^m) + \alpha x_c^N(p^m)$. From (29), for all $c \in \{1, 2\}$ and $\alpha \in [0, 1]$,

$$x_c(0) = x_c^B(p^m) \geq x_c(\alpha) \geq x_c^N(p^m) = x_c(1)$$

From (29), and by continuity of $\alpha \to x(\alpha)$, there exists $\overline{\alpha} \in (0, 1)$ such that for all $\alpha \in [0, \overline{\alpha})$ and all $c \in \{1, 2\}$, $x_c(\alpha) > x_c^B(p^{CEE})$, and for some $d \in \{1, 2\}$, $x_d(\overline{\alpha}) = x_d^B(p^{CEE})$. By construction of $\overline{\alpha}$, $x_{-d}(\overline{\alpha}) \geq x_{-d}^B(p^{CEE}) \geq x_{-d}^N(p^m)$. Note that $x_{-d} \to U_{-d}(p_{-d}^m, x_d, x_{-d})$ is concave with a maximum at $x_{-d} = x_{-d}^N(p_{-d}^m) = D_{-d}(p_{-d}^m)$. Therefore, the previous inequality implies that the median voter of country $-d$ prefers $x^B(p^{CEE})$ to $(x_d^B(p^{CEE}), x_{-d}(\overline{\alpha})) = x(\overline{\alpha})$. Moreover, from Property (9), $U_{-d}(p_{-d}^m, x(0)) > U_{-d}(p_{-d}^m, x(1))$. Since $U_{-d}(p_{-d}^m, x(\alpha))$ is strictly concave in α, the last inequality imply that she strictly prefers $x(\overline{\alpha})$ to $x(1) = x^N(p^m)$. By transitivity, she strictly prefers $x^B(p^{CEE})$ to

[20] To see the parallel with the sketch of the proof at the end of Section 4.3, note that the terms $\frac{\partial \Delta_1}{\partial p_1^r} + \frac{\partial \Delta_1}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}$ and $\frac{\partial \Delta_2}{\partial p_1^r} + \frac{\partial \Delta_2}{\partial x_1}\frac{\partial \hat{x}_1}{\partial p_1^r}$ correspond to the effect of p_1^r on $\Delta_1(p^r, \hat{x}(f, p^r))$ and $\Delta_2(p^r, \hat{x}(f, p^r))$, that is, on the gains from cooperation for representatives 1 and 2, respectively. As explained in the sketch of the proof, these effects determine the effect of p_1^r on B_1 and B_2, respectively, and thus on the terms of cooperation B_2/B_1. Note that the above terms differentiate $\Delta(p^r, \hat{x}(f, p^r))$ w.r.t. p_1^r keeping f fixed. So in the language of the sketch of the proof, we take the "as if" subsidies τ_1 and τ_2 fixed. Taking τ_1 and τ_2 fixed does not affect the sign of the effect of p_1^r on B_2/B_1 because we have shown that the denominator of the R.H.S. of (25), which collects all the indirect effects of p_1^r on τ_1 and τ_2, is always strictly positive.

$x^N(p^m)$. The proposition follows then from Lemma 1.□

Proof of Proposition 6. If D_c is unit elastic and D_{-c} is more convex than D_c, then we know from Proposition 3 that in any CEE p^{CEE}, $x_c^B(p^{CEE}) = x_c^N(p^{NEE})$ and $x_{-c}^B(p^{CEE}) > x_{-c}^N(p^{NEE})$. Since $G_{-c}{}' > 0$, this implies that the median voter in country c is strictly better off in the CEE than in the NEE. Since $U_{-c}(p_{-c}^m, x_c, x_{-c})$ is concave in x_{-c} with a maximum at $x_{-c} = D_{-c}(p_{-c}^m) = x_{-c}^N(p^{NEE})$, the median voter of country c is strictly worse off in the CEE than in the NEE.

A.5. Proofs for Section 4.5

Proof of Proposition 7. The cooperative electoral best response of each country c is given by the solution of $\max_{p_c^r>0} U_c(p_c^m, x^B(p^r))$. Observe that if we can prove that this program has a unique maximum, then existence of a CEE follows then from Brower's fixed point theorem, since from the maximum theorem, the continuity of $p_c^r \to U_c(p_c^m, x^B(p^r))$ implies the continuity of its unique maximum. To prove uniqueness of the cooperative electoral best response, it suffices to show that $p_c^r \to U_c(p_c^m, x^B(p^r))$ is strictly quasi concave.

For the isoelastic specification, for all $\varepsilon > 0$ such that $\varepsilon \neq 1$,

$$U_c(p_c^m, x_c^B(p^r)) = \frac{\varepsilon}{\varepsilon - 1}\left(\frac{p_c^r}{1 + \beta_c \frac{B_{-c}(p^r)}{B_c(p^r)}}\right)^{1-\varepsilon} + \beta_{-c}\frac{\varepsilon}{\varepsilon - 1}\left(\frac{p_{-c}^r}{1 + \beta_{-c}\frac{B_c(p^r)}{B_{-c}(p^r)}}\right)^{1-\varepsilon} - p_m^r\left(\frac{p_c^r}{1 + \beta_c\frac{B_{-c}(p^r)}{B_c(p^r)}}\right)^{-\varepsilon}.$$

From Lemma 3, the fraction $\frac{B_2(p^r)}{B_1(p^r)}$ in the above formula is the is the solution f to (17), which, in the case of the generalized Nash bargaining function,[21] yields

$$f = \frac{\pi_2 \Delta_1(p^r, \hat{x}(f, p^r))}{\pi_1 \Delta_2(p^r, \hat{x}(f, p^r))}, \tag{30}$$

where

$$\begin{cases} (\varepsilon - 1)\Delta_1(p^r, \hat{x}(f, p^r)) = [(1 + \beta_1 f)^{\varepsilon-1}(1 + (1 - \varepsilon)\beta_1 f) - 1](p_1^r)^{1-\varepsilon} + \varepsilon\beta_2\left[\left(1 + \frac{\beta_2}{f}\right)^{\varepsilon-1} - 1\right](p_2^r)^{1-\varepsilon} \\ (\varepsilon - 1)\Delta_2(p^r, \hat{x}(f, p^r)) = \left[\left(1 + \frac{\beta_2}{f}\right)^{\varepsilon-1}\left(1 + (1 - \varepsilon)\frac{\beta_2}{f}\right) - 1\right](p_2^r)^{1-\varepsilon} + \varepsilon\beta_1[(1 + \beta_1 f)^{\varepsilon-1} - 1](p_1^r)^{1-\varepsilon} \end{cases}. \tag{31}$$

Case $\varepsilon = 1$ (i.e., $G(x) = \ln(x)$ *and* $D(p) = 1/p$).

When $\varepsilon = 1$, $pD(p)\varepsilon(p)$ is constant, so from Lemma 5, $\frac{B_{-c}(p^r)}{B_c(p^r)}$ and $x_c^B(p^r)$ are independent of p_{-c}^r. Therefore, the induced utility function of the median voter can be written as

$$U_c(p_c^m, x^B(p^r)) = G_c(x_c^B(p_c^r)) + \beta_{-c}G_{-c}(x_{-c}^B(p_{-c}^r)) - p_c^m x_c^B(p_c^r).$$

From Lemma 4, $x_c^B(p_c^r)$ is strictly monotonic in p_c^r, and since $x_c \to G_c(x_c) - p_c^m x_c$ is strictly concave, $U_c(p_c^m, x^B(p^r))$ is strictly quasi concave in p_c^r, as needed.

Case $\varepsilon = 2$ (i.e., $G(x) = 2\sqrt{x}$ *and* $D(p) = 1/p^2$).

From (31), when $\varepsilon = 2$,

$$\begin{cases} \Delta_1(p^r, \hat{x}(f, p^r)) = -\frac{\beta_1^2 f^2}{p_1} + 2\frac{\beta_2^2}{f p_2} \\ \Delta_2(p^r, \hat{x}(f, p^r)) = -\frac{\beta_2^2}{f^2 p_2} + 2\frac{\beta_1^2 f}{p_1}, \end{cases}$$

and simple algebra shows that the solution to (30) is $f = \frac{\hat{\pi}_2\beta_2^{2/3}p_1^{1/3}}{\hat{\pi}_1\beta_1^{2/3}p_2^{1/3}}$, where $\hat{\pi}_c = (2\pi_c + \pi_{-c})^{1/3}$. Substituting the latter expression for f into $\hat{x}(f, p^r)$, we obtain $x_c^B(p) = D_c\left(\frac{p_c}{1 + \frac{\hat{\pi}_{-c}\beta_c^{1/3}\beta_{-c}^{2/3}p_c^{1/3}}{\hat{\pi}_c p_{-c}^{1/3}}}\right)$. Substituting $G(x) = 2\sqrt{x}$, $D(p) = 1/p^2$, and the latter expression for $x^B(p)$ into $U_1(p_1^m, x^B(p))$, we obtain

[21] Note that for this proof, we use the Nash bargainign function B. So this particular result may not be true for other bargaining functions.

$$U_1(p_1^m, x^B(p)) = 2\frac{1 + \frac{\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}p_1^{1/3}}{\hat{\pi}_1 p_2^{1/3}}}{p_1} + 2\beta_2\frac{1 + \frac{\hat{\pi}_1\beta_2^{1/3}\beta_1^{2/3}p_2^{1/3}}{\hat{\pi}_2 p_1^{1/3}}}{p_2} - p_1^m\left(\frac{1 + \frac{\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}p_1^{1/3}}{\hat{\pi}_1 p_2^{1/3}}}{p_1}\right)^2$$

$$= \frac{2}{p_1} + \frac{2\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}}{\hat{\pi}_1 p_2^{1/3}p_1^{2/3}} + \frac{2\beta_2}{p_2} + \frac{2\hat{\pi}_1\beta_2^{4/3}\beta_1^{2/3}}{\hat{\pi}_2 p_1^{1/3}p_2^{2/3}} - p_1^m\left(\frac{1}{p_1^2} + \frac{2\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}}{\hat{\pi}_1 p_2^{1/3}p_1^{5/3}} + \frac{\hat{\pi}_2^2\beta_1^{2/3}\beta_2^{4/3}}{\hat{\pi}_1^2 p_2^{2/3}p_1^{4/3}}\right).$$

Differentiating w.r.t. p_1, we get

$$\frac{\partial[U_1(p_1^m, x^B(p))]}{\partial p_1} = -\frac{2}{p_1^2} - \frac{4\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}}{3\hat{\pi}_1 p_2^{1/3}p_1^{5/3}} - \frac{2\hat{\pi}_1\beta_2^{4/3}\beta_1^{2/3}}{3\hat{\pi}_2 p_1^{4/3}p_2^{2/3}} + p_1^m\left(\frac{2}{p_1^3} + \frac{10\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}}{3\hat{\pi}_1 p_2^{1/3}p_1^{8/3}} + \frac{4\hat{\pi}_2^2\beta_1^{2/3}\beta_2^{4/3}}{3\hat{\pi}_1^2 p_2^{2/3}p_1^{7/3}}\right)$$

$$= \frac{1}{p_1^2}\left[-2 - \frac{4\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}p_1^{1/3}}{3\hat{\pi}_1 p_2^{1/3}} - \frac{2\hat{\pi}_1\beta_2^{4/3}\beta_1^{2/3}p_1^{2/3}}{3\hat{\pi}_2 p_2^{2/3}} + p_1^m\left(\frac{2}{p_1} + \frac{10\hat{\pi}_2\beta_1^{1/3}\beta_2^{2/3}}{3\hat{\pi}_1 p_2^{1/3}p_1^{2/3}} + \frac{4\hat{\pi}_2^2\beta_1^{2/3}\beta_2^{4/3}}{3\hat{\pi}_1^2 p_2^{2/3}p_1^{1/3}}\right)\right].$$

Each of the terms inside the bracket on the R.H.S. of the above equation is decreasing, some of them strictly decreasing, which proves that $U_1(p_1^m, x^B(p^r))$ is strictly quasi concave in p_1^r, as needed. The proof for country 2 is analogous.

Case $\varepsilon \to \infty$.

In what follows, we will show that for all $a, b > 0$ such that $a < b$, if we restrict voters in the electoral stage of the cooperative regime to elect a representative with a type in $[a, b]$, then for ε sufficiently large, such a restricted CEE exists, and it must tend to $p^{CEE} = p^m$ as $\varepsilon \to \infty$. Since we can choose a arbitrarily small and b arbitrarily large, this property implies the existence of an unrestricted CEE for ε sufficiently large, because as $\varepsilon \to \infty$, D_c becomes infinitely elastic, so given a strategy profile p^r close to p^m, electing type $p_c{}^r$ smaller than a or greater than b is not a profitable deviation for the median voter of country c. To prove existence of a restricted CEE, it suffices to show that for all $a, b > 0$ such that $a < b$, for ε sufficiently large, for all $p_{-c}^r \in [a, b]$, $p_c^r \to U_c(p_c^m, x^B(p^r))$ is strictly quasi concave on $[a, b]$. As shown in the proof of Proposition 2,

$$\frac{\partial[U_1(p_1^m, x_1^B(p^r))]}{\partial p_1^r} = \frac{\partial x_1^B}{\partial p_1^r}\left(\frac{p_1^r}{1 + \beta_1 f(p^r)} + \frac{\beta_2 p_2^r}{1 + \beta_2/f(p^r)}\frac{\frac{\partial x_2^B}{\partial p_1^r}}{\frac{\partial x_1^B}{\partial p_1^r}} - p_1^m\right), \tag{32}$$

and from Lemma 4, $\frac{\partial x_1^B}{\partial p_1^r} < 0$. So to prove the desired property, it suffices to show that for ε sufficiently large, for all $p_2^r \in [a, b]$, the term in parenthesis on the R.H.S. of (32) is strictly increasing in p_1^r on $[a, b]$.

From (31), for all $p^r \in \mathbb{R}_{++}^2$ and $f \in \mathbb{R}_{++}$, as $\varepsilon \to \infty$,

$$\frac{\Delta_1(p^r, \hat{x}(f, p^r))}{\Delta_2(p^r, \hat{x}(f, p^r))} = \frac{-\beta_1 f\left(\frac{p_1^r}{1 + \beta_1 f}\right)^{1-\varepsilon} + \beta_2\left(\frac{p_2^r}{1 + \frac{\beta_2}{f}}\right)^{1-\varepsilon} + o\left(\max\left\{\left(\frac{p_1^r}{1 + \beta_1 f}\right)^{1-\varepsilon}, \left(\frac{p_2^r}{1 + \frac{\beta_2}{f}}\right)^{1-\varepsilon}\right\}\right)}{-\frac{\beta_2}{f}\left(\frac{p_2^r}{1 + \frac{\beta_2}{f}}\right)^{1-\varepsilon} + \beta_1\left(\frac{p_1^r}{1 + \beta_1 f}\right)^{1-\varepsilon} + o\left(\max\left\{\left(\frac{p_1^r}{1 + \beta_1 f}\right)^{1-\varepsilon}, \left(\frac{p_2^r}{1 + \frac{\beta_2}{f}}\right)^{1-\varepsilon}\right\}\right)}$$

$$= \frac{-\beta_1 f R(\varepsilon, p^r, f) + \beta_2 + o(\max\{R(\varepsilon, p^r, f), 1\})}{-\frac{\beta_2}{f} + \beta_1 R(\varepsilon, p^r, f) + o(\max\{R(\varepsilon, p^r, f), 1\})}, \tag{33}$$

where $R(\varepsilon, p^r, f) \equiv \left(\frac{p_1^r}{1 + \beta_1 f}\frac{1 + \frac{\beta_2}{f}}{p_2^r}\right)^{1-\varepsilon}$. Let $f(\varepsilon, p^r)$ denote the solution to (30). We now show that as $\varepsilon \to +\infty$, $\frac{p_1^r}{1 + \beta_1 f(\varepsilon, p^r)}\frac{1 + \beta_2/f(\varepsilon, p^r)}{p_2^r}$ tends to 1 uniformly over all p^r. Suppose by contradiction that this is false. Then there exists a sequence p^n and $\varepsilon^n \to \infty$ such that $\frac{p_1^n}{1 + \beta_1 f(\varepsilon^n, p^n)}\frac{1 + \beta_2/f(\varepsilon^n, p^n)}{p_2^n}$ tends to some limit $l \in [0, +\infty]$ with $l \neq 1$. Suppose to fix ideas that $l < 1$, the proof in the case $l > 1$ is analogous. In that case, by definition of R, $R(\varepsilon^n, p^n, f(\varepsilon^n, p^n)) \to +\infty$, so from (33), $\frac{\Delta_1(p^n, \hat{x}(f(\varepsilon^n, p^n), p^n))}{\Delta_2(p^n, \hat{x}(f(\varepsilon^n, p^n), p^n))} \to -\lim f(\varepsilon^n, p^n) < 0$, which contradicts (30). Therefore, as $\varepsilon \to \infty$, the solution $f(\varepsilon, p^r)$ to (30) tends uniformly to a limit $f^\infty(p^r)$ which is the unique positive solution to

$$\frac{p_1^r}{1+\beta_1 f^\infty(p^r)} = \frac{p_2^r}{1+\dfrac{\beta_2}{f^\infty(p^r)}}. \tag{34}$$

As shown above, $R(\varepsilon, p^r, f(\varepsilon, p^r))$ must remain bounded as $\varepsilon \to \infty$. From (33), it must tend to some $R^\infty(p^r)$ uniformly over all p^r, where from (30), $R^\infty(p^r)$ is given by

$$f^\infty(p^r) = \frac{\pi_2(-\beta_1 f^\infty(p^r) R^\infty(p^r) + \beta_2)}{\pi_1\left(-\dfrac{\beta_2}{f^\infty(p^r)} + \beta_1 R^\infty(p^r)\right)} \Rightarrow R^\infty(p^r) = \frac{\beta_2}{\beta_1 f^\infty(p^r)}. \tag{35}$$

Moreover, by differentiating (30) w.r.t. p_1^r and then letting $\varepsilon \to \infty$, we obtain that $\frac{\partial f(\varepsilon, p^r)}{\partial p_1^r}$ converges uniformly over all p^r in the compact interval $[a, b]^2$. Therefore, $\frac{\partial f(\varepsilon, p^r)}{\partial p_1^r}$ must converge to $\frac{\partial f^\infty(p^r)}{\partial p_1^r}$ as $\varepsilon \to \infty$ (see, e.g., Rudin, 1976, Theorem 7.17). Differentiating (34), we obtain that

$$\frac{\partial\left[\dfrac{p_1^r}{1+\beta_1 f^\infty(p^r)}\right]}{\partial p_1^r} = \frac{\partial\left[\dfrac{p_2^r}{1+\beta_2/f^\infty(p^r)}\right]}{\partial p_1^r} > 0. \tag{36}$$

Using (7), (34), and (36), we obtain

$$\frac{\dfrac{\partial x_2^B}{\partial p_1^r}}{\dfrac{\partial x_1^B}{\partial p_1^r}} = \frac{\left(\dfrac{p_2^r}{1+\beta_2/f(\varepsilon, p^r)}\right)^{-\varepsilon-1}}{\left(\dfrac{p_1^r}{1+\beta_1 f(\varepsilon, p^r)}\right)^{-\varepsilon-1}} \frac{\dfrac{\partial\left[\dfrac{p_2^r}{1+\beta_2/f(\varepsilon, p^r)}\right]}{\partial p_1^r}}{\dfrac{\partial\left[\dfrac{p_1^r}{1+\beta_1 f(\varepsilon, p^r)}\right]}{\partial p_1^r}} \to_{\varepsilon\to\infty} \frac{1}{R^\infty(p^r)} = \frac{\beta_1 f^\infty(p^r)}{\beta_2},$$

so as $\varepsilon \to \infty$, the term in parenthesis on the R.H.S. of (32) tends towards

$$\frac{p_1^r}{1+\beta_1 f^\infty(p^r)} + \frac{\beta_2 p_2^r}{1+\beta_2/f^\infty(p^r)} \frac{1}{R^\infty(p^r)} - p_1^m = p_1^r - p_1^m. \tag{37}$$

The derivative of (37) w.r.t p_1^r is 1. By continuity, the derivative of the term in parenthesis on the R.H.S. of (32) converges uniformly over all p^r in the compact interval $[a, b]^2$, so it must converge uniformly to 1 (see, e.g., Rudin, 1976, Theorem 7.17). Thus, we have shown that for all bounded interval $[a, b]$, for ε sufficiently large, for all $p_2^r \in [a, b]$, $U_1(p_1^m, x_1^B(p^r))$ is strictly quasi concave in p_1^r on $[a, b]$. Moreover, we see from (37) that for all $[a, b]$, the electoral best response of the voters of country 1 to any p_2^r must tend to p_1^m as $\varepsilon \to \infty$, as needed.

Case $(\beta_1, \beta_2) \to (0, 0)$.

Clearly, as $\beta \to (0, 0)$, for the median voter of country c, electing an arbitrarily small or large type is not a profitable deviation from a strategy profile p^r close enough to p^m. Therefore, as argued in the case $\varepsilon \to \infty$, to prove the existence of a CEE, it suffices to show that for all compact intervals $[a, b]$, for β sufficiently small, for all $p_2^r \in [a, b]$, the term in parenthesis on the R.H.S. of (32) is strictly increasing in p_1^r on $[a, b]$, and that its root tends to p_1^m as $\beta \to (0, 0)$.

Note that as $(\beta_1, \beta_2) \to (0, 0)$, $\frac{\partial x_2^B}{\partial p_1^r}$ tends to 0, so the term in parenthesis on the R.H.S. of (32) tends towards $p_1^r - p_c^m$, whose derivative w.r.t. p_1^r is 1. The same continuity argument as in the proof of the case $\varepsilon \to \infty$ implies then that as $(\beta_1, \beta_2) \to (0, 0)$, the derivative of the term in parenthesis on the R.H.S. of (32) must tend to 1 uniformly over all $p^r \in [a, b]$, which shows that for all bounded interval $[a, b]$, for β sufficiently small, for all $p_2^r \in [a, b]$, $p_1^r \to U_1(p_1^m, x_1^B(p^r))$ is strictly quasi concave on $[a, b]$, as needed.□

A.6. Proofs in Section 5

Proof of Proposition 8. With utilitarian bargaining, $B_c = \pi_c$, so from (7), the outcome of the policy making stage of the cooperative regime is then $x_c^B(p^r) = D_c\left(\frac{p_c^r}{1+\beta_c \pi_{-c}/\pi_c}\right)$. Thus, the contribution of country c depends only on $p_c{}^r$, and the electoral stage of the cooperative regime is strategically equivalent to the game in which the voters of each country c control their own policy $x_c{}^B$, taking the policy of the other country x_{-c}^B as given. It is therefore strategically equivalent to the electoral stage of the noncooperative regime (see Section 4.1).□

A.7. Proofs in Section 6

The program that determines $x^B(p^r)$ and the transfer $t^B(p^r)$ is

$$\max_{x_1\geq 0, x_2\geq 0, t} [\pi_1 \ln(\Delta_1(p^r, x) + p_1^r t) + \pi_2 \ln(\Delta_2(p^r, x) - \alpha p_2^r t)]$$

where $\pi_1 + \pi_2 = 1$ and for all $c \in \{1, 2\}$,

$$\Delta_c(p^r, x) = G_c(x_c) + \beta_{-c} G_{-c}(x_{-c}) - p_c^r x_c - G_c(D_c(p_c^r)) - \beta_{-c} G_{-c}(D_{-c}(p_{-c}^r)) + p_c^r D_c(p_c^r).$$

The F.O.C. of this program are

$$\begin{cases} \pi_1 \dfrac{G'1(x_1) - p_1^r}{\Delta_1(p^r, x) + p_1^r t} + \pi_2 \dfrac{\beta_1 G'1(x_1)}{\Delta_2(p^r, x) - \alpha p_2^r t} = 0, \\ \pi_1 \dfrac{\beta_2 G'2(x_2)}{\Delta_1(p^r, x) + p_1^r t} + \pi_2 \dfrac{G'2(x_2) - p_2^r}{\Delta_2(p^r, x) - \alpha p_2^r t} = 0, \\ \dfrac{\pi_1 p_1^r}{\Delta_1(p^r, x) + p_1^r t} = \dfrac{\pi_2 \alpha p_2^r}{\Delta_2(p^r, x) - \alpha p_2^r t}. \end{cases}$$

Substituting the third condition into the first two, we obtain $G'1(x_1) = \frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}$ and $G'2(x_2) = \frac{p_1^r p_2^r}{p_1^r + \alpha \beta_2 p_2^r}$, and therefore

$$\begin{cases} x^B = \left(D_1\left(\dfrac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r} \right), D_2\left(\dfrac{p_1^r p_2^r}{p_1^r + \alpha \beta_2 p_2^r} \right) \right), \\ t^B = \pi_1 \dfrac{\Delta_2(p^r, x^B(p^r))}{\alpha p_2^r} - \pi_2 \dfrac{\Delta_1(p^r, x^B(p^r))}{p_1^r}, \end{cases}$$

which proves (11). The welfare of the median voter of country 1 for a given profile of representative p^r is then

$$\begin{aligned} V_1(p_1^m, x^B(p^r), t^B(p^r)) = G_1(x_1^B) + \beta_2 G_2(x_2^B) - p_1^m x_1^B + p_1^m t^B &= G_1(x_1^B)\left(1 + \frac{\pi_1 \beta_1 p_1^m}{\alpha p_2^r} - \frac{\pi_2 p_1^m}{p_1^r}\right) + G_1(D_1(p_1^r))\left(\frac{\pi_2 p_1^m}{p_1^r} - \frac{\pi_1 \beta_1 p_1^m}{\alpha p_2^r}\right) \\ &+ G_2(x_2^B)\left(\beta_2 + \frac{\pi_1 p_1^m}{\alpha p_2^r} - \frac{\pi_2 \beta_2 p_1^m}{p_1^r}\right) + G_2(D_2(p_2^r))\left(\frac{\pi_2 \beta_2 p_1^m}{p_1^r} - \frac{\pi_1 p_1^m}{\alpha p_2^r}\right) - x_1^B \frac{\pi_1 p_1^m}{\alpha} - x_2^B \frac{\pi_1 p_1^m}{\alpha} + D_2(p_2^r) \frac{\pi_1 p_1^m}{\alpha} \\ &- D_1(p_1^r) \pi_2 p_1^m. \end{aligned}$$

Differentiating w.r.t. p_1^r, using

$$\frac{\partial x_1^B}{\partial p_1^r} = \left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^2 D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right) \quad \text{and} \quad \frac{\partial x_2^B}{\partial p_1^r} = \alpha \beta_2 \left(\frac{p_2^r}{\alpha \beta_2 p_2^r + p_1^r}\right)^2 D'2\left(\frac{p_1^r p_2^r}{p_1^r + \alpha \beta_2 p_2^r}\right)$$

and regrouping the terms in factor of $D'1(p_1^r)$, $D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)$ and $D'2\left(\frac{p_1^r p_2^r}{\beta_2 p_2^r + p_1^r}\right)$, we obtain

$$\begin{aligned} \frac{\partial [V_1(p_1^m, x^B, t^B)]}{\partial p_1^r} &= -D'1(p_1^r) \frac{\beta_1 \pi_1 p_1^r p_1^m}{\alpha p_2^r} + D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3 (p_1^r - p_1^m) + D'2\left(\frac{p_1^r p_2^r}{p_1^r + \alpha \beta_2 p_2^r}\right)\left(\frac{p_2^r}{\alpha \beta_2 p_2^r + p_1^r}\right)^3 \alpha \\ &(\beta_2)^2 (p_1^r - p_1^m) + (G_1(x_1^B) - G_1(D_1(p_1^r))) \frac{\pi_2 p_1^m}{(p_1^r)^2} + (G_2(x_2^B) - G_2(D_2(p_2^r))) \frac{\pi_2 \beta_2 p_1^m}{(p_1^r)^2}. \end{aligned}$$

The F.O.C. of the electoral best response of country 1 is $\frac{\partial [V_1(p_1^m, x^B(p^r), t^B(p^r))]}{\partial p_1^r} = 0$. Using the above equation, this F.O.C. implies

$$\frac{p_1^r}{p_1^m} = \frac{\begin{aligned} &D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3 + D'2\left(\frac{p_1^r p_2^r}{p_1^r + \alpha \beta_2 p_2^r}\right)\left(\frac{p_2^r}{\alpha \beta_2 p_2^r + p_1^r}\right)^3 \alpha (\beta_2)^2 + D'1(p_1^r) \frac{\beta_1 \pi_1 p_1^r}{\alpha p_2^r} \\ &- \frac{\pi_2 (G_1(x_1^B(p^r)) - G_1(D_1(p_1^r)))}{(p_1^r)^2} - \frac{\pi_2 \beta_2 (G_2(x_2^B(p^r)) - G_2(D_2(p_2^r)))}{(p_1^r)^2} \end{aligned}}{D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3 + D'2\left(\frac{p_1^r p_2^r}{p_1^r + \alpha \beta_2 p_2^r}\right)\left(\frac{p_2^r}{\alpha \beta_2 p_2^r + p_1^r}\right)^3 \alpha (\beta_2)^2}. \tag{38}$$

Proof of Proposition 9. Suppose $D_1 = D_2 = D$, $\beta_1 = \beta_2 = 1$, $\alpha = 1$, $\pi_1 = \pi_2 = 1/2$, and $p_1^m = p_2^m = p^m$, and let (p^r, p^r) be a symmetric CEE. Then (38) implies

$$\frac{p^r}{p^m} = \frac{D'\left(\frac{p^r}{2}\right) + 2D'(p^r) + \frac{4}{(p^r)^2} \int_{\frac{p^r}{2}}^{p^r} pD'(p)dp}{D'\left(\frac{p^r}{2}\right)} = 2 + \frac{4 \int_{\frac{p^r}{2}}^{p^r} \left(p^r D'(p^r) + pD'(p) - \frac{p^r}{2} D'\left(\frac{p^r}{2}\right)\right) dp}{(p^r)^2 D'\left(\frac{p^r}{2}\right)} \tag{39}$$

It is straightforward to check that the equation $2^{-\varepsilon} - 1 + \frac{2^{1-\varepsilon}-1}{1-\varepsilon} = 0$ as a unique solution $\bar{\varepsilon}$. If we define $\overline{D}(p) \equiv p^{-\bar{\varepsilon}}$, then straightforward calculus shows that for $D = \overline{D}$, the integral on the R.H.S. of (39) is equal to 0 for any p^r. Therefore, if f is such that for all $p > 0$, $D(p) = f(\overline{D}(p))$, then $D'(p) = \overline{D}'(p)f'(\overline{D}(p))$ and

$$\frac{p^r}{p^m} = 2 + \frac{4\int_{\frac{p^r}{2}}^{p^r}\left(p^r\overline{D}'(p^r)f'(\overline{D}(p^r)) + p\overline{D}'(p)f'(\overline{D}(p)) - \frac{p^r}{2}\overline{D}'\left(\frac{p^r}{2}\right)f'\left(\overline{D}\left(\frac{p^r}{2}\right)\right)\right)dp}{(p^r)^2\overline{D}'\left(\frac{p^r}{2}\right)f'\left(\overline{D}\left(\frac{p^r}{2}\right)\right)}$$

$$= 2 + \frac{4\int_{\frac{p^r}{2}}^{p^r}\left(p^r\overline{D}'(p^r)\left[f'(\overline{D}(p^r]) - f'\left(\overline{D}\left(\frac{p^r}{2}\right)\right)\right) + p\overline{D}'(p)\left[f'(\overline{D}(p)) - f'\left(\overline{D}\left(\frac{p^r}{2}\right)\right)\right]\right)dp}{(p^r)^2\overline{D}'\left(\frac{p^r}{2}\right)f'\left(\overline{D}\left(\frac{p^r}{2}\right)\right)}$$

Since $x^B(p^r) = D(p^r/2)$, $x^B(p^r)$ is strictly greater (smaller) than $x^N(p^m)$ when $p^r/2$ is strictly smaller (greater) than p^m. Since $\widehat{D}' < 0$ and $f' > 0$, the above equation shows that this is the case when f' is strictly increasing (decreasing), or equivalently, when D is more convex than $\overline{D}$.□

Proof of Proposition 10. Suppose $\beta_2 = 0$ and let p^r be a CEE. Then the F.O.C. of the electoral best response of countries 1 (38) implies that

$$\frac{p_1^r}{p_1^m} = \frac{D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3 + D'1(p_1^r)\frac{\beta_1\pi_1 p_1^r}{\alpha p_2^r} - \left(G_1\left(D_1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\right) - G_1(D_1(p_1^r))\right)\frac{\pi_2}{(p_1^r)^2}}{D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3}.$$

From (11), $x_1^B(p^r) > x_1^N(p_2^m)$ if and only if $\frac{\alpha p_1^r p_2^r}{(\alpha p_2^r + \beta_1 p_1^r)p_1^m} < 1$, and from the above equation,

$$\frac{\alpha p_1^r p_2^r}{(\alpha p_2^r + \beta_1 p_1^r)p_1^m} = \frac{D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3 + D'1(p_1^r)\frac{\beta_1\pi_1 p_1^r}{\alpha p_2^r} - \left(G_1\left(D_1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\right) - G_1(D_1(p_1^r))\right)\frac{\pi_2}{(p_1^r)^2}}{D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^2}$$

$$= 1 + \frac{-D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^2\frac{\beta_1 p_1^r}{\alpha p_2^r + \beta_1 p_1^r} + D'1(p_1^r)\frac{\beta_1\pi_1 p_1^r}{\alpha p_2^r} + \frac{\pi_2}{(p_1^r)^2}\int_{\frac{1}{p_1^r}}^{\frac{\alpha p_2^r + \beta_1 p_1^r}{\alpha p_1^r p_2^r}}\frac{1}{t^3}D'1\left(\frac{1}{t}\right)dt}{D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^2},$$

where the second equality uses $\frac{\partial[G_1(D_1(1/t))]}{\partial t} = -D'1(1/t)/t^3$. Since $\pi_1 + \pi_2 = 1$, the above equation can be rewritten as follows:

$$\begin{aligned}
\left(\frac{\alpha p_1^r p_2^r}{(\alpha p_2^r + \beta_1 p_1^r)p_1^m} - 1\right)D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^2 &= \pi_1\left(D'1(p_1^r)\frac{\beta_1 p_1^r}{\alpha p_2^r} - D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^2\frac{\beta_1 p_1^r}{\alpha p_2^r + \beta_1 p_1^r}\right) \\
&+ \pi_2\left(\frac{1}{(p_1^r)^2}\int_{\frac{1}{p_1^r}}^{\frac{\alpha p_2^r + \beta_1 p_1^r}{\alpha p_1^r p_2^r}}\frac{1}{t^3}D'1\left(\frac{1}{t}\right)dt\right. \\
&\left. - D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^2\frac{\beta_1 p_1^r}{\alpha p_2^r + \beta_1 p_1^r}\right) \\
&= \frac{\pi_1\beta_1}{\alpha p_2^r(p_1^r)^2}\left(D'1(p_1^r)(p_1^r)^3 - D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3\right) \\
&+ \frac{\pi_2}{(p_1^r)^2}\left(\int_{\frac{1}{p_1^r}}^{\frac{\alpha p_2^r + \beta_1 p_1^r}{\alpha p_1^r p_2^r}}\left(\frac{1}{t^3}D'1\left(\frac{1}{t}\right) - D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3\right)dt\right).
\end{aligned}$$

The above equation shows that $\frac{\alpha p_1^r p_2^r}{(\alpha p_2^r + \beta_1 p_1^r)p_1^m}$ is smaller (greater) than 1, and thus $x_1^B(p^r)$ is greater (smaller) than $x_1^N(p_2^m)$ when $p \to p^3 D'(p)$ is increasing (increasing), or equivalently, when D is more convex than $1/p^2$.

As for country 2, the F.O.C. of its electoral best response of countries 2 can be obtained by substituting $\beta_1 = 0$ in (38), inverting the roles of 1 and 2, and replacing α by $1/a$, which yields

$$\frac{p_2^r}{p_2^m} = \frac{D'2(p_2^r) + D'1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\left(\frac{\alpha p_1^r}{\alpha p_2^r + \beta_1 p_1^r}\right)^3 \frac{(\beta_1)^2}{\alpha} - \left(G_1\left(D_1\left(\frac{\alpha p_1^r p_2^r}{\alpha p_2^r + \beta_1 p_1^r}\right)\right) - G_1(D_1(p_1^r))\right)\frac{\pi_1 \beta_1}{(p_2^r)^2}}{D'1(p_1^r) + D'2\left(\frac{p_1^r p_2^r}{p_1^r + \alpha\beta_2 p_2^r}\right)\left(\frac{p_2^r}{\alpha\beta_2 p_2^r + p_1^r}\right)^3 \alpha(\beta_2)^2}.$$

The above equation shows that $p_2^r > p_2^m$, which from (11), implies that $x_2^B(p^r) < x_2^N(p_2^m)$.□

A.8. Proofs in Section 7.1

Using the notations introduced in Section 7, since $g_c = x_c + \beta_{-c} x_{-c}$, we have that $x_c = \frac{g_c - \beta_{-c} g_{-c}}{1 - \beta_1\beta_2}$, so the utility function of representative $p_c{}^r$ in (13) can be rewritten as a function of the vector of public good levels (g_1, g_2) as follows: for all $g \in \mathbb{R}_+^2$,

$$V_c(p_c^r, g) \equiv U_c\left(p_c^r, \left(\frac{g_1 - \beta_2 g_2}{1 - \beta_1\beta_2}, \frac{g_2 - \beta_1 g_1}{1 - \beta_1\beta_2}\right)\right) = G_c(g_c) - p_c^r \frac{g_c}{1 - \beta_1\beta_2} + p_c^r \frac{\beta_{-c} g_{-c}}{1 - \beta_1\beta_2}. \tag{40}$$

Proof of Proposition 11. For a given x_{-c}, the best response of representative $p_c{}^r$ to x_{-c} is the solution to the program $\max_{x_c \geq 0} U_c(p_c^r, x_c, x_{-c})$. The F.O.C. of that convex program is $x_c = \max\{0, D_c(p_c^r) - \beta_{-c} x_{-c}\}$. Since these best responses are continuous, there exists an equilibrium, and any equilibrium x can be of one and only one of the following three types:

Type 1: $D_1(p_1^r) - \beta_2 x_2 > 0$ *and* $D_2(p_2^r) - \beta_1 x_1 > 0$.

In that case,

$$\begin{cases} x_1 = D_1(p_1^r) - \beta_2 x_2 \\ x_2 = D_2(p_2^r) - \beta_1 x_1 \end{cases} \Rightarrow \begin{cases} x_1 = \dfrac{D_1(p_1^r) - \beta_2 D_2(p_2^r)}{1 - \beta_1\beta_2} \\ x_2 = \dfrac{D_2(p_2^r) - \beta_1 D_1(p_1^r)}{1 - \beta_1\beta_2} \end{cases}.$$

The two inequalities $D_1(p_1^r) - \beta_2 x_2 > 0$ and $D_2(p_2^r) - \beta_1 x_1 > 0$ are satisfied if and only if $D_1(p_1^r)/D_2(p_2^r) \in (\beta_2, 1/\beta_1)$. The corresponding level of public good is $g_c = x_c + \beta_{-c} x_{-c} = D_c(p_c^r)$.

Type 2: $D_1(p_1^r) - \beta_2 x_2 \leq 0$ *and* $D_2(p_2^r) - \beta_1 x_1 > 0$.

In that case, $x_1 = 0$ and $x_2 = D_2(p_2^r)$. The inequality $D_2(p_2^r) - \beta_1 x_1 > 0$ is clearly satisfied, and $D_1(p_1^r) - \beta_2 x_2 \leq 0$ is satisfied when $D_1(p_1^r)/D_2(p_2^r) \leq \beta_2$.

Type 3: $D_1(p_1^r) - \beta_2 x_2 > 0$ *and* $D_2(p_2^r) - \beta_1 x_1 \leq 0$.

In that case $x_1 = D_1(p_1^r)$ and $x_2 = 0$. The inequality $D_1(p_1^r) - \beta_2 x_2 > 0$ is clearly satisfied, and $D_2(p_2^r) - \beta_1 x_1 \leq 0$ is satisfied when $D_1(p_1^r)/D_2(p_2^r) \geq \frac{1}{\beta_1}$.□

Proof of Proposition 12. For all p^r such that $D_1(p_1^r)/D_2(p_2^r) \in (\beta_2, 1/\beta_1)$, if $p_c{}^r$ is a noncooperative electoral best response to p_{-c}^r, it must be a solution of $\max_{p_c^r > 0} U_c(p_c^m, x^N(p^r))$, so it satisfies the F.O.C. $\partial[U_1(p_1^m, x^N(p^r))]/\partial p_1^r = 0$. Using (14) and the fact that for all $p \in \mathbb{R}_{++} G'_c(D_c(p)) = p$, we obtain

$$\frac{\partial[U_c(p_c^m, x^N(p^r))]}{\partial p_c^r} = D'c(p_c^r)\left(p_c^r - \frac{p_c^m}{1 - \beta_1\beta_2}\right).$$

Substituting the above expression into the F.O.C. $\frac{\partial[U_1(p_1^m, x^N(p^r))]}{\partial p_1^r} = 0$, we obtain $p_1^{NEE} = \frac{p_1^m}{1 - \beta_1\beta_2}$. From Proposition 11, $g_1^{NEE} = D_1\left(\frac{p_1^m}{1 - \beta_1\beta_2}\right)$.

As argued in Section 7.1, the outcome of the NEE is the outcome of the game in which the median voter of each country c controls g_c, taking g_{-c} as given, and her payoff is $V_c(p_c^m, g)$ where V_c is defined in (40). Because of the presence of positive externalities captured by the term $\frac{\beta_{-c} g_{-c}}{1 - \beta_1\beta_2}$ in (40) it is obvious that the Nash equilibrium of that game is inefficient, so there exists g^* that is strictly preferred by both median voters. From Lemma 1, g^* is strictly preferred by a majority of voters in both countries as well. To conclude the proof, it remains to show that any such vector g^* must be such that $g_1^* > D_1\left(\frac{p_1^m}{1 - \beta_1\beta_2}\right)$ and $g_2^* > D_2\left(\frac{p_2^m}{1 - \beta_1\beta_2}\right)$. Suppose that one of these inequality is violated, say the latter for concreteness. Then

$$V_1(p_1^m, g^*) < V_1\left(p_1^m, g_1^*, D_2\left(\frac{p_2^m}{1 - \beta_1\beta_2}\right)\right) \leq \max_{g_1 \geq 0} V_1\left(p_1^m, g_1, D_2\left(\frac{p_2^m}{1 - \beta_1\beta_2}\right)\right) = V_1\left(p_1^m, D_1\left(\frac{p_1^m}{1 - \beta_1\beta_2}\right), D_2\left(\frac{p_2^m}{1 - \beta_1\beta_2}\right)\right),$$

so the median voter of country 1 does not strictly prefer g^* to g^{NEE}, a contradiction.

A.9. Proofs in Section 7.2

In the policy making stage of the cooperative regime, the bargaining process maximizes $B(\Delta(p^r, x))$. To simplify the algebra, it will be convenient to rescale the gains from cooperation $\Delta_c(p^r, x)$ by the multiplicative factor $1/p^r_c$, and to express it as a function of g instead of x. That is, with a slight abuse of notations, we define

$$\Delta_c(p^r, g) \equiv \frac{1}{p^r_c}(V_c(p^r_c, g) - V_c(p^r_c, g^N(p^r))).$$

Since the generalized Nash bargaining solution that is, $B(\Delta) = \Delta_1^{\pi_1}\Delta_2^{\pi_2}$ is invariant to linear transformation of the payoffs, this rescaling is inconsequential.[22] Using (40) and (14), we can write:

$$\Delta_c(p^r, g) = \frac{G_c(g_c)}{p^r_c} - \frac{g_c}{1-\beta_1\beta_2} + \frac{\beta_{-c}g_{-c}}{1-\beta_1\beta_2} - \frac{G_c(D_c(p^r_c))}{p^r_c} + \frac{D_c(p^r_c)}{1-\beta_1\beta_2} - \frac{\beta_{-c}D_{-c}(p^r_{-c})}{1-\beta_1\beta_2}. \tag{41}$$

The program (6) can then be equivalently rewritten as a function of g as follows:

$$\max_{\substack{g_1 \ge \beta_2 g_2 \\ g_2 \ge \beta_1 g_1}} B(\Delta(p^r, g)). \tag{42}$$

The following Lemma establishes inter alia Eq. (15).

Lemma 6. *If we denote* $\hat{g}(f, p^r) \equiv \left(D_1\left(\frac{1-\beta_1 f}{1-\beta_1\beta_2}p^r_1\right), D_2\left(\frac{1-\beta_2/f}{1-\beta_1\beta_2}p^r_2\right)\right)$, *then the equation*

$$f = \frac{\frac{\partial B}{\partial \Delta_2}(\Delta(p^r, \hat{g}(f, p^r)))}{\frac{\partial B}{\partial \Delta_1}(\Delta(p^r, \hat{g}(f, p^r)))} \tag{43}$$

has a unique solution f^*, *and* $g^B(p^r) = \hat{g}(f^*, p^r)$.

Proof. One can see from (41) that $\Delta_c(p^r, g)$ is strictly concave in g, so the same arguments as in Lemma 2 imply that the solution $g^B(p^r)$ to the program (42) is unique. To see why it must also be interior, suppose by contradiction that it is not. Then $g^B_c(p^r) = \beta_{-c}g^B_{-c}(p^r)$ for some c, so $x^B_c(p^r) = 0$ and

$$U_{-c}(p^r_{-c}, x^B(p^r)) = G_{-c}(x^B_{-c}(p^r)) - p^r_{-c}x^B_{-c}(p^r) \le \max_{x_{-c}} G_{-c}(x_{-c}) - p^r_{-c}x_{-c} \le G_{-c}(D_{-c}(p^r_{-c})) - p^r_{-c}D_{-c}(p^r_{-c}) \le U_{-c}(p^r_{-c}, x^N(p^r)).$$

The above inequality contradicts property (9).

Since $g^B(p^r)$ is an interior solution to (42), it must satisfy the F.O.C. $\partial[B(\Delta(p^r, g))]/\partial g = 0$. Using (41), these conditions can be rewritten as

$$\begin{cases} \frac{\partial B}{\partial \Delta_1}(\Delta(p^r, x))\left(\frac{G'1(g_1)}{p^r_1} - \frac{1}{1-\beta_1\beta_2}\right) + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r, x))\frac{\beta_1}{1-\beta_1\beta_2} = 0 \\ \frac{\partial B}{\partial \Delta_1}(\Delta(p^r, x))\frac{\beta_2}{1-\beta_1\beta_2} + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r, x))\left(\frac{G'2(g_2)}{p^r_2} - \frac{1}{1-\beta_1\beta_2}\right) = 0 \end{cases}$$

The above system can be viewed as a linear system in $(G'1(g_1), G'2(g_2))$ whose solution is

$$G'c(g_c) = \left(1 - \beta_c\frac{\frac{\partial B}{\partial \Delta_{-c}}(\Delta(p^r, g))}{\frac{\partial B}{\partial \Delta_c}(\Delta(p^r, g))}\right)\frac{p^r_c}{1-\beta_1\beta_2}.$$

Applying the function D_c and using the notation B_c introduced in Section 4.2, we obtain

$$g^B_c(p^r) = D_c\left(\frac{1-\beta_c B_{-c}/B_c}{1-\beta_1\beta_2}p^r_c\right). \tag{44}$$

The latter equation can be rewritten as $g^B(p^r) = \hat{g}(f, p^r)$, where f is a solution to (43).

Reciprocally, let f^* be a solution to (43), and let $g^* = \hat{g}(f^*, p^r)$. Then using successively the definition of $\hat{g}$ and (43), we obtain

[22] With other bargaining functions B, the results might depend on the particular affine transformation of utility functions one uses in (6) to compute $x^B(p^r)$. However, for the affine transformation $U_c(p^r_c, x)/p^r_c$ used in (41), the reader can check that, as in the basic model, our proofs hold more generally for any function $B(\Delta)$ that satisfies (9) and (10).

$$\frac{\partial B}{\partial \Delta_1}(\Delta(p^r, g^*))\left(\frac{G'1(g_1^*)}{p_1^r} - \frac{1}{1-\beta_1\beta_2}\right) + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r, g^*))\frac{\beta_1}{1-\beta_1\beta_2} = \frac{\partial B}{\partial \Delta_1}(\Delta(p^r, g^*))\left(\frac{1-\beta_1 f^*}{1-\beta_1\beta_2} - \frac{1}{1-\beta_1\beta_2}\right) + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r, g^*))\frac{\beta_1}{1-\beta_1\beta_2} = -\frac{\partial B}{\partial \Delta_1}(\Delta(p^r, g^*))\frac{\beta_1 f^*}{1-\beta_1\beta_2} + \frac{\partial B}{\partial \Delta_2}(\Delta(p^r, g^*))\frac{\beta_1}{1-\beta_1\beta_2} = 0,$$

which shows that g^* satisfies the F.O.C. of the program (42) for c=1, as derived at the beginning of this proof. An analogous reasoning shows that it also satisfies this F.O.C. for c=2. Therefore, g^* is a solution to (42), so $g^* = g^B(p^r)$.

Thus, we have shown that the solutions to (43) are in a one to one relationship with the solutions of the program (42). Since the latter program has a unique solution, (43) has a unique solution too.□

Lemma 7. *For all* p^r, $g_c^B(p^r) \geq g_c^N(p^r)$.

Proof. We first show that $x_c^B(p^r) \geq x_c^N(p^r)$. Suppose this is not the case, then

$$U_{-c}(p_{-c}^r, x^B) < U_{-c}(p_{-c}^r, x_c^N, x_{-c}^B) \leq \max_{x_{-c}\geq 0} U_{-c}(p_{-c}^r, x_c^N, x_{-c}) = U_{-c}(p_{-c}^r, x^N),$$

which violates property (9). Since both countries make greater contributions, the level of public goods in both countries must be higher, as needed.

The next lemma is the equivalent of Lemma 4 in the basic model. It shows that in the cooperative regime, the level of public good g_c^B in each country is strictly decreasing in the type of its representative p_c^r, and thus that by choosing p_c^r, the voters of each country effectively choose their level of public good (although p_c^r also affects the level of public good in the other country g_{-c}^B, see Lemma 9 below).

Lemma 8. *For all* p^r *such that* $\frac{D_1(p_1^r)}{D_2(p_2^r)} \in \left(\beta_2, \frac{1}{\beta_1}\right)$, $\frac{\partial g_c^B}{p_c^r} < 0$.

Proof. The proof of this lemma follows the same logic as the proof of Lemma 4. Suppose by contradiction that there exists $\hat{p}^r$ such that $\frac{\partial g_c^B}{p_c^r}(\hat{p}^r) \geq 0$. From (44),

$$\frac{\partial g_c^B}{\partial p_c^r} = \left(\frac{1-\beta_c B_{-c}/B_c}{1-\beta_1\beta_2} - \frac{\beta_c p_c^r}{1-\beta_1\beta_2}\frac{\partial[B_{-c}/B_c]}{\partial p_c^r}\right) D'c\left(\frac{1-\beta_c B_{-c}/B_c}{1-\beta_1\beta_2}p_c^r\right).$$

By assumption, $D'c < 0$ and $\frac{\partial g_c^B}{p_c^r}(\hat{p}^r) \geq 0$. Substituting these two inequalities into the above equation, we obtain

$$\frac{\partial[B_{-c}/B_c]}{\partial p_c^r}(\hat{p}^r) > 0. \tag{45}$$

Together with (15), (45) implies that $\frac{\partial g_{-c}^B}{\partial p_c^r}(\hat{p}^r) < 0$. Using (40), (15) and $D_c = (G_{,c})^{-1}$, simple calculus yields

$$\frac{\partial[V_c(\hat{p}_c^r, g^B(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r} = \frac{(1-\beta_c B_{-c}/B_c)p_c^r}{1-\beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} - \frac{\hat{p}_c^r}{1-\beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} + \frac{\beta_{-c}\hat{p}_c^r}{1-\beta_1\beta_2}\frac{\partial g_{-c}^B}{\partial p_c^r} - \frac{\beta_c(B_{-c}/B_c)p_c^r}{1-\beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} - \frac{(\hat{p}_c^r - p_c^r)}{1-\beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} + \frac{\beta_{-c}\hat{p}_c^r}{1-\beta_1\beta_2}\frac{\partial g_{-c}^B}{\partial p_c^r},$$

where the functions in the above expressions are evaluated at $(p_c^r, \hat{p}_{-c}^r)$. If we set $p_c^r = \hat{p}_c^r$ and substitute $\frac{\partial g_c^B}{p_c^r}(\hat{p}^r) \geq 0$ and $\frac{\partial g_{-c}^B}{\partial p_c^r}(\hat{p}^r) < 0$ in the above expression, we obtain that $\frac{\partial[V_c(\hat{p}_c^r, g^B(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r}(\hat{p}_c^r) < 0$. Since $g^B(\hat{p}^r)$ is Pareto optimal for the representatives $\hat{p}^r$, the above inequality implies that $\frac{\partial[U_{-c}(\hat{p}_{-c}^r, g^B(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r}(\hat{p}_c^r) \geq 0$. Using (14), we always have $\frac{\partial[V_{-c}(\hat{p}_{-c}^r, g^N(p_c^r, \hat{p}_{-c}^r))]}{\partial p_c^r}(\hat{p}^r) < 0$. The last two inequalities imply that the gain from cooperation for representative $\hat{p}_{-c}^r$ increases as p_c^r increases around $\hat{p}_c^r$. Using the notation $\Delta(p^r, g)$, this means that

$$\frac{\partial[\Delta_{-c}(p^r, g^B(p^r))]}{\partial p_c^r}(\hat{p}^r) > 0. \tag{46}$$

Differentiating (41) and using $D_c = (G_{,c})^{-1}$, we obtain

$$\frac{\partial[\Delta_c(p^r, x^B(p^r))]}{\partial p_c^r} = \frac{\frac{(1-\beta_c B_{-c}/B_c)(p_c^r)^2}{1-\beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} - G(g^B)}{(p_c^r)^2} - \frac{\frac{\partial g_c^B}{\partial p_c^r} - \beta_{-c}\frac{\partial g_{-c}^B}{\partial p_c^r}}{1-\beta_1\beta_2} - \frac{(p_c^r)^2\frac{\partial g_c^N}{\partial p_c^r} - G(g^N)}{(p_c^r)^2} + \frac{\frac{\partial g_c^N}{\partial p_c^r}}{1-\beta_1\beta_2}$$

$$= -\frac{\beta_c B_{-c}/B_c}{1-\beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} + \frac{\beta_{-c}}{1-\beta_1\beta_2}\frac{\partial g_{-c}^B}{\partial p_c^r} + \frac{\beta_1\beta_2}{1-\beta_1\beta_2}\frac{\partial g_c^N}{\partial p_c^r} - \frac{G(g^B) - G(g^N)}{(p_c^r)^2}.$$

Substituting the inequalities $\frac{\partial g_c^B}{\partial p_c^r}(\hat{p}^r) \geq 0$, $\frac{\partial g_{-c}^B}{\partial p_c^r}(\hat{p}^r) < 0$, $\frac{\partial g_c^N}{\partial p_c^r}(p_c^r) = D'_c(p_c^r) < 0$, and $G(g^B(p^r)) \geq G(g^N(p^r))$ (the latter inequality follows from Lemma 7) into the above equation, we obtain that $\frac{\partial[\Delta_c(p^r, x^B(p^r))]}{\partial p_c^r}(\hat{p}^r) < 0$. Using property (10), the latter inequality and (46) imply then that B_{-c}/B_c must be strictly increasing in $p_c{}^r$ at $p^r = \hat{p}^r$, a contradiction with (45).□

Proof of Proposition 13. From (40), for any p^r such that $\frac{D(p_1^r)}{D(p_2^r)} \in \left(\beta_1, \frac{1}{\beta_2}\right)$, the welfare of the median voter in the cooperative policy equilibrium is

$$V_c(p_c^m, g^B(p^r)) = G_c(g_c^B(p^r)) - p_c^m \frac{g_c^B(p^r) - \beta_{-c} g_{-c}^B(p^r)}{1 - \beta_1\beta_2}.$$

Differentiating w.r.t. p_c^r, and using (15) and $D_c = (G'_c)^{-1}$, we obtain

$$\frac{\partial[V_c(p_c^m, g^B(p^r))]}{\partial p_c^r} = \frac{(1 - \beta_c B_{-c}/B_c)p_c^r}{1 - \beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} - \frac{p_c^m}{1 - \beta_1\beta_2}\frac{\partial g_c^B}{\partial p_c^r} + \frac{p_c^m \beta_{-c}}{1 - \beta_1\beta_2}\frac{\partial g_{-c}^B}{\partial p_c^r}.$$

Suppose that p^r is such that $D(p_1^r)/D(p_2^r) \in (\beta_1, 1/\beta_2)$ and $p_c{}^r$ is a cooperative electoral best response to p_{-c}^r. Then $p_c{}^r$ must satisfy the F.O.C. $\partial[V_1(p_1^m, g^B(p^r))]/\partial p_1^r = 0$. Substituting the above equation into the F.O.C. and using the fact that from Lemma 8, $\partial g_c^B/\partial p_c^r \neq 0$, we obtain

$$(1 - \beta_c B_{-c}/B_c)p_c^r = p_c^m\left(1 - \beta_{-c}\frac{\partial g_{-c}^B}{\partial p_c^r}/\frac{\partial g_c^B}{\partial p_c^r}\right).$$

Substituting the above equation into (15), we obtain

$$g_c^B(p^r) = D_c\left(\frac{p_c^m}{1 - \beta_1\beta_2}\left(1 - \beta_{-c}\frac{\partial g_{-c}^B}{\partial p_c^r}/\frac{\partial g_c^B}{\partial p_c^r}\right)\right).$$

From Proposition 12, if $\tilde{p}_c^r$ is a noncooperative electoral best response of country c to some $\tilde{p}_{-c}^r$, then $g_c^N(\tilde{p}^r) = D_c\left(\frac{p_c^m}{1 - \beta_1\beta_2}\right)$. By comparing the latter two equations, we see that $g_c^B(p^r) - g_c^N(\tilde{p}^r)$ has the same sign as $\frac{\partial g_{-c}^B}{\partial p_c^r}/\frac{\partial g_c^B}{\partial p_c^r}$, which from Lemma 8 has the opposite sign of $\partial g_{-c}^B/\partial p_c^r$, which from (15) has the opposite sign of $\partial[B_{-c}/B_c]/\partial p_c^r$. Thus, $g_c^B(p^r) - g_c^N(\tilde{p}^r)$ has the same sign as $\partial[B_{-c}/B_c]/\partial p_c^r$, as needed.□

A.10. Proofs in Section 7.3

The next lemma is the equivalent of Lemma 5 in the basic model: it shows that the curvature of D_c determines how the type of a representative $p_c{}^r$ affects B_{-c}/B_c.

Lemma 9. *Let φ_c be such that for all $p_c^r > 0$, $D_c(p_c^r) = \varphi_c(\ln(1/p_c^r))$. If φ_c is convex (concave), then for all p^r such that $\frac{D_1(p_1^r)}{D_2(p_2^r)} \in \left(\beta_2, \frac{1}{\beta_1}\right)$, $\frac{\partial[B_{-c}/B_c]}{\partial p_c^r} > 0\ (< 0)$, where B_{-c} and B_c are defined in (8).*

Proof. This proof follows the same steps as the proof of Lemma 5. We prove it for c=1, the proof for c=2 is analogous. From Lemma 6, B_2/B_1 is given by the solution f to the fixed point condition (43) for c=1. Therefore, it suffices to show that if φ_1 is convex (concave), then $\partial f/\partial p_1^r > 0\ (< 0)$.

Differentiating (43) w.r.t. p_1^r, and following the same algebraic steps as in Lemma 5, we obtain that $\partial f/\partial p_1^r$ is given by (25), where $\hat{x}$ is replaced by $\hat{g}$ (see Lemma 6) and Δ is given by (41). Since $D'_c < 0$, by definition of $\hat{g}$ (see Lemma 6), we have $\partial\hat{g}_1/\partial f > 0$ and $\partial\hat{g}_2/\partial f < 0$. Moreover, from (41), $\frac{\partial\Delta_1}{\partial g_1}(p_1^r, \hat{g}_1) = \frac{G'1(g_1)}{p_1^r} - \frac{1}{1-\beta_1\beta_2} = -\frac{\beta_1 f}{1-\beta_1\beta_2} < 0$, and $\frac{\partial\Delta_1}{\partial g_2}(p_1^r, \hat{g}_1) = \frac{\beta_2}{1-\beta_1\beta_2} > 0$. Therefore, the term $\frac{\partial\Delta_1}{\partial g_1}\frac{\partial\hat{g}_1}{\partial f} + \frac{\partial\Delta_1}{\partial g_2}\frac{\partial\hat{g}_2}{\partial f}$ in the denominator on the R.H.S. of (25) is negative. An analogous reasoning implies that $\frac{\partial\Delta_2}{\partial g_1}\frac{\partial\hat{g}_1}{\partial f} + \frac{\partial\Delta_2}{\partial g_2}\frac{\partial\hat{g}_2}{\partial f}$ is positive. Thus, $\frac{\partial\Delta_1}{\partial g_1}\frac{\partial\hat{g}_1}{\partial f} + \frac{\partial\Delta_1}{\partial g_2}\frac{\partial\hat{g}_2}{\partial f}$ and $\frac{\partial\Delta_1}{\partial g_1}\frac{\partial\hat{g}_1}{\partial f} + \frac{\partial\Delta_1}{\partial g_2}\frac{\partial\hat{g}_2}{\partial f}$ have the same sign as in Lemma 5, so the same reasoning implies that $\frac{\partial f}{\partial p_1^r}$ is positive (negative) when $\frac{\partial\Delta_1}{\partial p_1^r} + \frac{\partial\Delta_1}{\partial g_1}\frac{\partial\hat{g}_1}{\partial p_1^r}$ and $-\left(\frac{\partial\Delta_2}{\partial p_1^r} + \frac{\partial\Delta_2}{\partial g_1}\frac{\partial\hat{g}_1}{\partial p_1^r}\right)$ are both positive (negative). Therefore, to complete the proof, it suffices to show that this is the case when φ_1 is strictly convex (concave).

By definition of φ_1, for all $p_1 > 0$, $p_1 D'_1(p_1) = -\varphi'_1(\ln(1/p_1))$. Using the latter equation, (41) and the definition of $\hat{g}_1$ in Lemma 6, we obtain

$$-\left(\frac{\partial\Delta_2}{\partial p_1^r} + \frac{\partial\Delta_2}{\partial g_1}\frac{\partial\hat{g}_1}{\partial p_1^r}\right) = \frac{\beta_1 D'1(p_1^r)}{1 - \beta_1\beta_2} - \frac{(1 - \beta_1 f)\beta_1}{(1 - \beta_1\beta_2)^2}D'1\left(\frac{(1 - \beta_1 f)p_1^r}{1 - \beta_1\beta_2}\right) = \frac{\beta_1}{(1 - \beta_1\beta_2)p_1^r}\left[\varphi'1\left(\ln\left(\frac{1 - \beta_1\beta_2}{(1 - \beta_1 f)p_1^r}\right)\right) - \varphi'1\left(\ln\left(\frac{1}{p_1^r}\right)\right)\right], \tag{47}$$

and

$$\frac{\partial\Delta_1}{\partial p_1^r}+\frac{\partial\Delta_1}{\partial g_1}\frac{\partial\hat{g}_1}{\partial p_1^r} = -\frac{G_1(g_1)}{(p_1^r)^2}+\frac{G_1(D_1(p_1^r))}{(p_1^r)^2} - D'1(p_1^r) + \frac{D'1(p_1^r)}{1-\beta_1\beta_2} \quad + \left(\frac{G'c(g_c)}{p_c^r} - \frac{1}{1-\beta_1\beta_2}\right)\frac{1-\beta_1 f}{1-\beta_1\beta_2}D'1\left(\frac{(1-\beta_1 f)p_1^r}{1-\beta_1\beta_2}\right)$$

$$= \frac{\beta_1}{(1-\beta_1\beta_2)p_1^r}\left[\begin{array}{l} \frac{1-\beta_1\beta_2}{\beta_1 p_1^r}\left(G_1(D_1(p_1^r)) - G_1\left(D_1\left(\frac{(1-\beta_1 f)p_1^r}{1-\beta_1\beta_2}\right)\right)\right)\\ -\beta_2\varphi'1\left(\ln\left(\frac{1}{p_1^r}\right)\right)+f\varphi'1\left(\ln\left(\frac{1-\beta_1\beta_2}{(1-\beta_1 f)p_1^r}\right)\right)\end{array}\right]$$

Simple calculus yields $\frac{\partial[G_1(D_1(p))]}{\partial p} = pD'1(p) = -\varphi'1(\ln(1/p))$, so the intermediate value theorem implies that there exists $p \in \left(\frac{(1-\beta_1 f)p_1^r}{1-\beta_1\beta_2}, p_1^r\right)$ such that

$$G_1(D(p_1^r)) - G_1\left(D_1\left(\frac{(1-\beta_1 f)p_1^r}{1-\beta_1\beta_2}\right)\right) = -\left(p_1^r - \frac{(1-\beta_1 f)p_1^r}{1-\beta_1\beta_2}\right)\varphi'1\left(\ln\left(\frac{1}{p}\right)\right) = \frac{\beta_1 p_1^r(\beta_2 - f)}{1-\beta_1\beta_2}\varphi'1\left(\ln\left(\frac{1}{p}\right)\right).$$

Substituting the above expression into the above expression for $\frac{\partial\Delta_1}{\partial p_1^r}+\frac{\partial\Delta_1}{\partial g_1}\frac{\partial\hat{g}_1}{\partial p_1^r}$, we obtain

$$\frac{\partial\Delta_1}{\partial p_1^r}+\frac{\partial\Delta_1}{\partial g_1}\frac{\partial\hat{g}_1}{\partial p_1^r} = \frac{\beta_1}{(1-\beta_1\beta_2)p_1^r} \quad \times\left[\beta_2\left(\varphi_{,1}\left(\ln\left(\frac{1}{p}\right)\right) - \varphi_{,1}\left(\ln\left(\frac{1}{p_1^r}\right)\right)\right) + f\left\{\varphi_{,1}\left(\ln\left(\frac{1-\beta_1\beta_2}{(1-\beta_1 f)p_1^r}\right)\right) - \varphi_{,1}\left(\ln\left(\frac{1}{p}\right)\right)\right\}\right] \tag{48}$$

From Lemma 7, $D_1\left(\frac{(1-\beta_1 f)p_1^r}{1-\beta_1\beta_2}\right) > D_1(p_1^r)$, so $\frac{(1-\beta_1 f)p_1^r}{1-\beta_1\beta_2} < p < p_1^r$, and $\ln\left(\frac{1}{p_1^r}\right) < \ln\left(\frac{1}{p}\right) < \ln\left(\frac{1-\beta_1\beta_2}{(1-\beta_1 f)p_1^r}\right)$. Therefore, (47) and (48) are both positive (negative) when $\varphi_{,1}$ is positive (negative), as needed.□

The proof of Proposition 14 follows immediately from the following proposition, and from Definitions 2 and 3.

Proposition 17. *Let p^r and $\tilde{p}^r$ be two profiles of representatives such that $D_1(p_1^r)/D_2(p_2^r) \in (\beta_2, 1/\beta_1)$, $D_1(\tilde{p}_1^r)/D_2(\tilde{p}_2^r) \in (\beta_2, 1/\beta_1)$, p_c^r is a cooperative electoral best response to p_{-c}^r, and $\tilde{p}_c^r$ is a noncooperative electoral best response to $\tilde{p}_{-c}^r$. If $D_c(p)$ is more (less) convex than $\ln(1/p)$, then $g_c^B(p^r)$ is greater (smaller) than $g_c^N(\tilde{p}^r)$.*

Proof. From Proposition 13, $x_c^B(p^r)$ is greater (smaller) than $x_c^N(\tilde{p}^r)$ if $\partial[B_{-c}/B_c]/\partial p_c^r$ is positive (negative). Lemma 9 imply that this is the case when $D_c(p)$ is more (less) convex than $\ln(1/p)$, as needed.□

References

Acemoglu, Daron, 2003. Why not a political coase theorem? Social conflict, commitment, and politics. J. Comput. Econ. 31, 620 652.
Alesina, Alberto, 1988. Credibility and policy convergence in a two-party system with rational voters. Am. Econ. Rev. 78, 796 805.
Auten, Gerald E., Sieg, Holger, Clotfelter, Charles T., 2002. Charitable giving, income, and taxes: an analysis of panel data. Am. Econ. Rev. 92 (1), 371 382.
Besley, Tim, Coate, Stephen, 2003. Centralized versus decentralized provision of local public goods: a political economy approach. J. Public Econ. 87, 2611 2637.
Brennan, Geoffrey, Buchanan, James, 1980. The Power to Tax: Analytical Foundations of a Fiscal Constitution. Cambridge University Press, New York, NY.
Brooks, Arthur C., 2007. Income tax policy and charitable giving. J. Policy Anal. Manag. 26 (3), 599 612.
Brückner, Matthias, 2000. Unanimity Versus Consensus Bargaining Games: Strategic Delegation and Social Optimality. University of Bonn, Mimeo.
Buchanan, James, Faith, R., 1987. Secession and the limits of taxation. Am. Econ. Rev. 77, 1023 1031.
Buchholz, Wolfgang, Haupt, Alexander, Peters, Wolfgang, 2005. International environmental agreements and strategic voting. Scandivanian J. Econ. 107 (1), 175 195.
Chari, V.V., Jones, Larry E., Marimon, Ramon, 1997. The economics of split-ticket voting in representative democracies. Am. Econ. Rev. 87 (5), 957 976.
Cheikbossian, Guillaume, 2000. Federalism, distributive politics and representative democracy. Econ. Gov. 2, 105 122.
Christiansen, Nels, 2013. Strategic delegation in a legislative bargaining model with pork and public goods. J. Public Econ. 97, 217 229.
Coase, Ronald, 1960. The problem of social cost. J. Law Econ. 3 (1), 1 44.
Cornes, Richard, 1993. Dyke maintenance and other stories: some neglected types of public goods. Q. J. Econ. 108, 259 271.
DelRossi, Alison F., Inman, Robert P., 1999. Changing the price of pork: the impact of local cost sharing on legislators'demands for distributive public goods. J. Public Econ. 71, 247 273.
Dur, Robert, Roelfsema, Hein, 2005. Why does centralisation fail to internalise policy externalities? Public Choice 122 (3/4), 395 416.
Feldstein, Martin S., 1975. The income tax and charitable contributions: part II the impact of religious, education, and other organizations. Natl. Tax. J. 27, 209 226.
Gradstein, Mark, 2004. Political bargaining in a federation: Buchanan meets coase. Eur. Econ. Rev. 48, 983 999.
Grandmont, Jean Michel, 1978. Intermediate preferences and the majority rule. Econometrica 46, 317 330.
Harstad, Bård, 2007. Harmonization and side payments in political cooperation. Am. Econ. Rev. 97 (3), 871 889.
Harstad, Bård, 2008. Do side payments help? Collective decisions and strategic delegation. J. Eur. Econ. Assoc. 6 (2 3), 468 477.
Harstad, Bård, 2010. Strategic delegation and voting rules. J. Public Econ. 94, 102 113.
Inman, Robert P., Rubinfeld, Daniel L., 1997. Rethinking federalism. J. Econ. Perspect. 11, 43 64.
Kehoe, Patrick J., 1989. Policy cooperation among benevolent governments may be undesirable. Rev. Econ. Stud. 56 (2), 289 296.
Kempf, Hubert, Rossignol, Stéphane, 2013. National politics and international agreements. J. Public Econ. 100, 93 105.
Loeper, Antoine, 2011. Coordination in heterogeneous federal systems. J. Public Econ. 95, 900 912.
Loeper, Antoine, 2015. Public Good Provision and Cooperation among Representative Democracies. Working Paper.
Mailath, G.J., Postlewaite, A., 1990. Asymmetric information bargaining problems with many agents. Rev. Econ. Stud. 57 (3), 351 367.
North, Douglass C., 1990. A transaction cost theory of politics. J. Theor. Polit. 2 (4), 355 367.
Oates, Wallace E., 1972. Fiscal Federalism. Harcourt-Brace, New York.
Olson, Mancur, 1965. The Logic of Collective Action. Harvard University Press, Cambridge, MA.
Persson, Torsten, Tabellini, Guido, 1992. The politics of 1992: fiscal policy and european integration. Rev. Econ. Stud. 59 (4), 689 701.

Persson, Torsten, Tabellini, Guido, 1995. Double-edged incentives: institutions and policy coordination. Handb. Int. Econ. 3, 1973–2030, (Chapter 38).
Redoano, Michela, Scharf, Kimberly, 2004. The political economy of policy centralization: direct versus representative democracy. J. Public Econ. 88 (3–4), 799–817.
Rogoff, Kenneth, 1985. Can international monetary policy cooperation be counterproductive? J. Int. Econ. 18, 199–217.
Rudin, Walter, 1976. Principles of Mathematical Analysis Third Edition. McGraw-Hill.
Sandler, Todd, 1998. Global and regional public goods: a prognosis for collective action. Fisc. Stud. 19 (3), 221–247.
Schelling, Thomas C., 1960. The Strategy of Conflict. Harvard University Press, Cambridge, MA.
Segendorff, Björn, 1998. Delegation and threat in bargaining. Games Econ. Behav. 23 (2), 266–283.
Tabellini, Guido, 1990. Domestic politics and the international coordination of fiscal policies. J. Int. Econ. 28 (3/4), 245–265.
Wildasin, David E., 1987. Distortionary taxation and demand estimation for public goods. Reg. Sci. Urban Econ. 19, 353–379.
Williamson, Oliver, 1985. The Economic Institutions of Capitalism. Free Press, New York, NY.